DOCUMENT RESUME

ED 027 625 EA 002 065

Big City Title I Evaluation Conference (Fort Lauderdale, Florida, December 27-30, 1966), 1967 Report.
Pittsburgh Public Schools, Pa.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Elementary and Secondary Education.
Pub Date 67
Grant-OEG-1-7-000059-0059
Note-409p.
EDRS Price MF-$1.75 HC-$20.55
Descriptors-Case Studies (Education), *Disadvantaged Youth, *Evaluation Methods, *Federal Programs, *Metropolitan Areas, Models, *Program Evaluation
Identifiers-*ESEA Title I

This conference report presents the position taken by large cities in evaluating projects under ESEA Title I. Since this was the first educational legislation in history to call for evaluation at the local level of programs generated by Federal funds, most districts had little time to design project evaluation and to obtain adequate research staffs, facilities, and basic data. The conference was held to aid in perfecting project evaluation procedures in order to gain continuing congressional support for the Title I program. Organized around a theme of sharing evaluation assumptions and operational procedures, the report is divided into three major sections: Principles of evaluation, evaluation based on theoretical models, and evaluation case studies. Section I includes an overview of ESEA evaluation and a summary of evaluation principles, while section II discusses the theoretical bases of the Columbus and Pittsburgh evaluation programs. Section III presents case studies for Cincinnati, Dade County, Detroit, Los Angeles, Milwaukee, New York, Oakland, and Philadelphia. (TT)

ED027625

BIG CITY TITLE I
EVALUATION
CONFERENCE

1967 REPORT

EA 002 065

PITTSBURGH PUBLIC SCHOOLS
SIDNEY P. MARLAND, JR., SUPERINTENDENT

ED027625

BIG CITY TITLE I EVALUATION CONFERENCE

1967 REPORT

THE BOARD OF PUBLIC EDUCATION · SCHOOL DISTRICT OF PITTSBURGH, PENNSYLVANIA

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Pittsburgh Public Schools

Sidney P. Marland, Jr., Superintendent

The research reported herein was performed pursuant to contract OEC1-7-000059-0059 with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

LIST OF CONFERENCE PARTICIPANTS

FROM THE BIG CITIES

William Ashbaugh, Milwaukee

Alden Badal, Oakland

Joan Bollenbacher, Cincinnati

Minerva S. Desing, Philadelphia

Mary Jane Duda, Pittsburgh

Ralph Hall, Dade County

Robert S. Lankton, Detroit

Norma W. Nickens, Washington

Malcolm M. Provus, Pittsburgh

Frank Toggenburger, Los Angeles

J. Wayne Wrightstone, New York

CONSULTANTS

Richard A. Dershimer, Executive Secretary, AERA, Washington, D. C.

Egon Guba, Director, National Institute for the Study of Educational Change, Indiana University

James M. Lipham, Chairman, Department of Educational Administration, University of Wisconsin

Daniel Stufflebeam, Director, Evaluation Center, Ohio State University

O. E. OBSERVERS

C. Butler, Office of Education

Frieda Denemark, Programs Branch, Office of Education

Stuart Westerlund, Consultant, Bureau of Elementary and Secondary Education, Office of Education

TABLE OF CONTENTS

LIST OF CONFERENCE PARTICIPANTS iii

INTRODUCTION vii

CHAPTER 1 PRINCIPLES OF EVALUATION 1

I An Overview of ESEA Evaluation 3
Malcolm M. Provus

II Summary of Evaluation Principles 13
Conference Participants

CHAPTER 2 EVALUATION BASED ON THEORETICAL MODELS 27

I The Columbus Evaluation Program 29
Michael D. Hock

II The Pittsburgh Evaluation Program 101
Mary Jane Duda

CHAPTER 3 A COMPENDIUM OF EVALUATION CASE STUDIES 261

I Cincinnati 263
Submitted by Joan Bollenbacher

II Dade County 277
Submitted by Ralph Hall

III Detroit 301
Submitted by Robert S. Lankton

IV Los Angeles 327
Submitted by Frank Toggenburger

TABLE OF CONTENTS (contd.)

V Milwaukee 359
Submitted by William Ashbaugh

VI New York 375
Submitted by J. Wayne Wrightstone

VII Oakland 391
Submitted by Alden W. Badal

VIII Philadelphia 445
Submitted by Minerva S. Desing
and
Franklin Institutes Research Laboratories

Introduction

Title I of the Elementary and Secondary Education Act passed by Congress in 1965 was the first educational legislation in history to call for an evaluation at the local level to determine the effectiveness of programs generated by federal funds. As such, it created tremendous demands on the recipients of these funds.

Title I was put into effect almost overnight with the intent of compensating for the limited experience of poor, deprived, and handicapped children. Most districts had little or no time to plan their ESEA projects, much less to design project evaluations. Research staffs and facilities were inadequate, and staff added subsequently did not have the advantage of being involved from the start. Basic data suddenly required were not available. Much had to be done to define the scope of the evaluation task, to invent useful data gathering instruments, and to secure information on short notice from those who were deeply engaged in project operations. Yet, evidence of project activity and effectiveness and evaluations capable of providing such evidence were needed in order to gain congressional support for the fiscal year 1967-1968. It was no doubt with a sense of urgency that Office of Education decision makers searched for existent evaluation designs and strategies.

The position taken by the big cities on ESEA evaluation was of special importance to the Office of Education. Most of these cities operated with a considerable degree of independence from state and federal guidelines. The big cities serve a relatively large number of underprivileged students, and most possess research staffs presumably capable of supporting a serious interpretation of the evaluation clause of ESEA. Therefore, a knowledge of big-city evaluation activity was particularly important to O.E. It was for this reason that the Pittsburgh Evaluation Model Validation Project was amended to include a conference for discussion among research directors of some of the big cities. Accordingly, a conference was held in Fort Lauderdale, Florida, December 27-30, 1966 with selected representatives from the big cities, consultants, and O. E. observors in attendance.

The organizing theme of the conference was the sharing of evaluation assumptions and operational procedures. All parties to the conference were surprised at their agreement on certain evaluation purposes although great variation in evaluation methodology obtained. On the third and final day of the conference all participants agreed as to the worth of a position paper which would attempt to summarize and explicate those areas of agreement as well as document samples of evaluation work in a number of big-city systems in an attempt to better understand local evaluation strategy and intent. Evaluation work in two of the big-city systems involved the development and use of evaluation models.

This publication resulted from the decisions described above and is, accordingly, divided into three major areas: principles of evaluation, evaluation based on theoretical models, and evaluation case studies.

Chapter 1

PRINCIPLES OF EVALUATION

Part I

AN OVERVIEW OF ESEA EVALUATION

It is apparent that Title I evaluation can serve at least three major purposes: quality control, cost benefit analysis, and improved decision making. Quality control requires the establishment of procedures to monitor and modify programs so as to insure uniform products which meet acceptable standards. Cost benefit analysis examines the relationship between cost of a program and the value of its products relative to the cost of other products of similar value. Finally, improved decision making is possible when new sources of information about either benefits, costs, program quality, or program feasibility are at hand.

The last of these purposes of evaluation appears to be most widely recognized to date. Yet, there seems to be general agreement in the Office of Education that current information about Title I programs is so minimal as to preclude even a description of Title I program activity in most American schools. Information as to benefits, costs, and program operations is simply not available. The possibility of comparing educational programs on a cost effectiveness basis is still but a Defense Department dream in the minds of a few Office of Education executives.

The first purpose, the desire to maintain quality control standards for Title I projects through evaluation apparatus, appears to be uppermost in the minds of a good many Office of Education staffers as well as the cause of continual effort and frustration. Generally, Office of Education personnel have sought to insure program quality by establishing categorical restrictions on fund use and by urging that more sophisticated evaluation designs be incorporated in program proposals.

This latter approach to evaluation assumes that school districts will be more likely to attend to program quality if they know that program products will be rigorously appraised. This strategy necessarily emphasizes the definition of program outputs in quantified terms suitable for use in experimental designs. In this strategy the government has been abetted by "research types" in universities. And herein lies an unrecognized paradox: as state and federal government officers seek to encourage program quality by means of rigorous product assessment, they may thwart quality control work at the only point it will be effective--the local level. This local work is usually of the process assessment type in which evaluators systematically collect and weigh data descriptive of ongoing program activity.

The desirability of submitting both Title I products and processes to evaluation so as to guide policy makers at the national level will find universal support. However, it is our belief that process evaluation for program improvement must occur at the local level before national assessments of Title I can be meaningful.

All Title I programs are new programs in the sense that they are new to the school district and to most of the personnel involved. These programs are in a "becoming stage" for staff. Procedures must change with experience. And as procedures change, as the possible and impossible are sorted out, goals must change too. However, when goals have been fixed, when students have been frozen into experimental and control groups to satisfy rigorous evaluation designs which assume stable treatments, the dynamics of essential program change are constricted out of existence.

Education may be defined as the art of accurately predicting changes in human behavior through applied social science. Because methodology is uncertain and conditions are obscure, predictions are often inaccurate. When dealing with underprivileged students, about whom we know even less than the average school client, our predictions are notoriously unreliable. It follows that educational programs devised for the underprivileged are unreliable. Programs with the same name can differ more from class to class than do programs with different names. Upon inspection, Miss Brown's remedial reading program may turn out to be a group therapy class. Miss Smith's therapy class may turn out to be a high school psychology lecture series.

Clearly, the task defined by Title I legislation is to increase education program payloads for the underprivileged and to prove it. Before this is possible, however, we will have to improve our program development procedures and know that we have improved them. It is important to note that improvement in procedure is not always immediately reflected in performance. For example, in aviation a change in wing design may be ineffective until coupled with increased horsepower. So in education a change in instructional material may be inconsequential until coupled with a new mix of students or a new teaching technique.

We must be able to demonstrate that procedures exist for the improvement of programs resulting in educational benefit to the underprivileged. Some program development work in universities, regional labs, and centers goes forward. However, it is obvious that unless massive aid is given to research in program development work, little change in practice will be seen in this decade. That federal policy reflects this reality can be inferred from the emphasis being given to Title III of ESEA. Title III is the "tinker and try" title of ESEA. However, the $75,000,000 available for Title III in 1966-1967 runs the risk of dilution because it is being spread across the entire spectrum of the American public schools population in contrast to the over a billion dollar Title I program funds directed toward those with greatest need. Furthermore, O.E.'s present policy appears to be to encourage Title III dissemination projects over developmental projects. The assumption that something of value to disseminate exists may in many cases not be warranted because so few programs have been properly developed.

Ultimately, programs will improve only if teachers, administrators, and students in most of America's classrooms become involved in a comprehensive effort to review and improve their work. Such an effort requires a careful study of present procedures, a detailed analysis of program events and their sequence, and the designing of a series of small experiments to test the value of program components.

Shades of Action Research? Maybe. But there is considerable evidence in both industry and education that only when personnel responsible for conducting a program are involved in its examination and revision will the program improve and endure. To improve educational programs we must use Title I to give school personnel a sense of freedom to admit error, to revise programs, and to creatively risk failure secure in the belief that continuous program evaluation will eventually provide success.

In this light, the fact that there has been little opportunity for carefully planned Title I programs is not as serious as some have claimed. Programs can be planned after they have begun. "In midstream" program planning has the advantage of providing staff with information based on pragmatic consequence rather than speculation. That is, classrooms become labs for culling program development information. Further, it becomes possible for an entire program staff to contribute to a policy decision and thereby feel a sense of responsibility for its consequences. These benefits, of course, ensue only when "in midstream" program planning is built in to a local Title I effort--that is, when evaluation is interpreted to mean staff self-evaluation and continuous program assessment.

The broadening of the superintendent's decision-making base to include program staff decisions may be viewed by some as a loss of administrative power. On the contrary, one of the most important considerations in an administrator's policy-making deliberations is the amount of staff support a decision will command. For this reason, superintendents often establish committees or councils as a sounding board for policy prior to enactment. One great disadvantage of such sounding boards is that they must react to a hypothetical issue without benefit of total staff reaction to "the real thing." When a superintendent can involve his entire staff in actual program revision and policy-making experience, his decisions as the chief school officer are likely to be realistic and enduring.

When Title I evaluation is used as a mechanism for the development and improvement of school programs, it takes on the appearance of a staff training strategy. Staff activity includes systematic study of program variables such as student entry behavior, student-teacher interaction, pupil performance on interim tasks, and student and teacher indices of attitude and satisfaction. Evaluation becomes a vehicle for training staff to meet the changing demands of project activity. Such training will eventually serve all information needs for process and product evaluation. It also provides continuous data on staff, pupils, and program for management decision making. It reveals new independent variables and stabilizes program variables, both of which are essential to eventual program benefit analysis. Judgment about program benefit can be based on product assessment only when a program is stable--when it continues to be what staff and administration think it is. The products of stable programs can be compared only when all of the factors which can change the product, other than program, have been identified and taken into account. Therefore, to expect to evaluate a program on the basis of its products before evaluating its processes will generally lead to frustration and failure.

Yet, process and product evaluation need not be incompatible in the same evaluation project. Only planning of event sequence and patience are needed. Program development as a function of "in midstream" planning and training must come first. Experimental design work must be deferred. To expect significant shifts in criterion scores before the final form of a program has evolved, is to expect a caterpillar to fly.

When program is defined by staff as a changing dynamic which must undergo developmental stages of growth, it becomes possible to understand, measure, and place a value on any program at any time without compulsive reference to standardized achievement test scores or other quantified criteria. Scores are not to be discounted in either the short term or long term, but they must lose their sacredness as the criterion of program effectiveness. Only when applied to the last stages of program development are they meaningful, and then only when they are part of an experimental design which makes provision for many complex factors.

Some officials in the Office of Education are obviously aware of the need to de-emphasize short-term product assessment so that internal local evaluation processes may be employed to strengthen programs. Yet pressure from the Congress and elsewhere may be too great to resist. A similar pressure exists locally from political action groups which in their desire to see tangible student benefits do themselves a disservice by ignoring the complex, maturational nature of any new school program.

Part II

SUMMARY OF EVALUATION PRINCIPLES

Distinction Between Goals and Objectives

Good evaluation starts with a clear definition of goals, or long-range educational aims. The goals of the educational enterprise are the projected results of the individual programs which constitute the entire enterprise. As such, goals are determined by the educational institution in the context of community, society, and projected future needs.

Objectives are the intermediate ends which lead to the realization of goals. Thus they are of greater concern to the evaluation team than are goals since the objectives provide the bases for the development of programs as well as methods used to measure program effects. While the value of program objectives is not a concern of the evaluation team, the evaluation process may be used to clarify and define objectives, and thereby increase their relevance and practicality.

Information Collection

Generally speaking, information collection includes the time scheduling of information requirements, the selection and/or development of appropriate instruments, a determination of sampling methods to be used, the administration of instruments under standard conditions, and related activities. Flexibility of methods, techniques, and research design must be maintained so that they are appropriate for each of the programs evaluated and for the particular stage at which information is collected for each program.

The collection of information descriptive of an ongoing program is an important early step in evaluation. The characteristics of students entering a program, their competencies, attitudes, needs, and interests, the special activities of students, teachers, and administrators participating in the program--all these must be determined, and the information should be quantified if possible.

Generally, adequate instruments do not exist for the description and recording of student, teacher, and administrator behavior. Most school districts must either distort phenomena to fit standardized commercial instruments or must use ad hoc questionnaires without evidence as to their reliability. In either case, a gross distortion of fact is likely to result. Even the reporting of simple attendance figures may be subject to errors, especially if these data are collected and summarized by a person who is more interested in artistry than in statistics. Close attention must be given to the creation of forms for supplying data and to the directions accompanying these forms. If accuracy is desired, the collection and tabulation of statistical data should be an established procedure controlled by the evaluator. Summary reports prepared by someone else may not stand actual scrutiny.

Information Reporting

Once information has been collected, the evaluator is responsible for the compilation of data and the production and editing of evaluation reports, including the necessary tables and graphs. The report format may be a joint undertaking of the evaluation and operations unit. Complete evaluation reports or appropriate summaries should be distributed to all participating district personnel, district administrators, board members, state and federal agencies, and other interested groups. The feedback of evaluation information to the field staff, particularly to teachers, is of vital importance.

These diverse groups may require reports written for different purposes. The Superintendent and Board may have interests in an evaluation quite different from those of a program staff. Public special interest groups may be concerned only with the success of special groups of children, and they may narrowly define success as dramatic changes in performance on standardized tests. Students themselves may apply very different but reasonable criteria to the evaluation program: their own interests, enthusiasm, and attitude toward school. The parents of these children may have similar evaluation interests.

Regardless of the purpose of the report, the timing of dissemination is crucial. Evaluation reports, perfect though they may be, submitted one year or even one day after a program has been refunded are of little value to teachers and administrators and of little interest to anyone else. A perceptive but unsophisticated report that does not follow the canons of academic research may be subject to scholarly criticism, but is of more value when submitted in time to effect changes in an ongoing program than is a technically perfect report submitted too late. Quality in written reports is highly desirable, but the report must also be made available in time to influence persons party to or interested in the public schools.

Elements of Support Necessary to Evaluation

Climate

Probably more important than any other single factor to the success of evaluation is that there be a climate which encourages investigation throughout the school system. Many evaluation efforts fail, or succeed only minimally, because teachers and administrators lack understanding of, and therefore commitment to, the role of research in education. Instruction in the rationale for evaluation, research procedures, data processing, and data usage must be provided in such a way as to truly involve all sectors of the school system in the evaluative effort. Such in-service training would help to keep alive a conscious awareness of common purpose; such awareness, unfortunately, often becomes submerged in the day-to-day operations of diverse programs.

A climate which encourages inquiry can make evaluation a continuing activity and an integral part of the educational process. Teachers should be encouraged to raise critical questions and then search for answers in systematic ways. The more active teachers are in evaluation, the more likely pupils are to evaluate their own work. Only if the spirit of inquiry is established and maintained will teachers, administrators, and pupils be convinced of the value of filling out another time-consuming questionnaire. Certainly, the validity of evaluative data is open to serious question if they have to be gathered by forced intervention.

Administrative Philosophy and Structure

Basic to a good climate for thoughtful inquiry in education is an administrative philosophy that is committed to evaluation. Today continuing evaluation of established programs is needed, as well as evaluation of projects that are experimental or innovative in nature. Procedures, methods, facilities, and learning materials need to be regarded as means to an end, and subject to continual improvement. It may well be that existing programs that were good in other times or under different conditions no longer yield the same benefits for pupils they once did. More enlightened doubt about present practices will lead to more convictions and more innovations.

Also embodied in this philosophy is the disposition to establish two-way communication with concerned public and professional personnel;

to recognize and use findings of previously completed research of other cities and other agencies; and to cooperate in common research projects of other large cities.

There are a variety of school district administrative arrangements that lend themselves to the assumption of the responsibilities attendant to the development, operation, and evaluation of educational programs. A school district may use its existing administrative organization, it may establish a separate evaluation unit, or it may contract with an outside agency.

Whatever the structure decided upon, the research director should be organizationally placed so that he is able to interact freely and frequently with associate superintendents or other line and staff personnel who are able to mobilize the resources of the school district both within and across organizational substructures. He should also have immediate access to the general superintendent of schools or persons equally empowered to render final judgments. Moreover, these relationships should be characterized by a high degree of mutual professional trust.

The evaluation unit itself should not be subordinate to those in charge of the program being evaluated. Evaluation personnel should serve in a staff or advisory capacity to project personnel rather than reporting "in line" to them. Such organizational placement is essential in order for independent, valid evaluations to be obtained.

Budget

Administrative philosophy and structure play an important role in insuring that adequate funds are supplied for evaluation activity, although just what adequate budget is cannot be prescribed. Some cities are presently using about 3 percent of total ESEA Title I funds for evaluation purposes. Certainly 5 percent of such funds could be considered reasonable. In fact, evaluation on a broad scale, applying to regular school programs as well as to innovative ones, would be greatly enhanced if that percentage of the school system's total operational budget were scheduled to research--including the associated aspects of evaluation, development, and demonstration.

Regardless of the amount, an important budget consideration in research is internal flexibility. There need to be opportunities to make changes within a total allocation to provide funds for special contingencies--contract services, in-service training, new learning materials, unusual equipment, and developmental work for the exploration of new ideas.

Personnel

Adequate budget, in part, provides for adequate personnel. A crucial need exists for well-trained research and evaluation personnel who are senstive to the sociological, psychological, political, and economic demands of our society so that they may be of maximum help in context assessment; who are technically competent in research, statistics, and measurement to assist with input assessment; who are administratively skilled in working interpersonally with a wide variety of general and technical personnel to assist in process assessment; and who are professionally sensitive to the problems of the schools in our present society to assist with product evaluation and policy making.

In view of this great need, those persons already within a school system, who know it well from practical experience, need to be encouraged to develop their research interests through further training with the knowledge that there will be opportunities for leadership in research and evaluation in their school systems. Present research staff should be up-graded through internship programs. Persons trained in academic research who are brought into the schools must be trained again--in the practical use of their theory under noncontrolled conditions. Formal course work in methodology of research and statistical techniques is a necessary aid to evaluation, but university credits and degrees do not make a research worker or an evaluator unless they are combined with the ability to apply these skills in a public school setting.

Facilities

Not only does adequate budget provide for personnel, it must also include provisions for facilities. With an operation as extensive and complex as evaluation has become, many things are required: offices, office machines of all types, warehouses and storage space, conference rooms, and space for experimental utilization of equipment materials and methods. Since these kinds of facilities are usually available in one form or another, attention must now be turned to a more crucial present need--improved school system facilities and capacities for information management.

Retrieval of information for analysis is becoming more difficult as the volume of relevant data increases. Special data processing systems should be planned and made operational for exclusive use in managing information related to instruction. Such systems must be sophisticated enough to employ efficient methods for data collection and use. Random access capabilities are required so as to be able to draw from vast data banks of pupil, teacher, and school system information.

Effective evaluation and communication are dependent upon a competent technology. Methods of data retrieval and storage currently in use are hopelessly antiquated. For example, much information is presently available on all students, including testing results, teacher evaluations, counselor records, extracurricular interests, and family histories. Yet because such data are recorded on so many forms and stored in so many different places, they are difficult, if not impossible to retrieve. Computer technology is absolutely essential for the utilization of all information sources. Only through an adequate data processing system can all potential and existing information be collected, synthesized, made immediately accessible, and utilized to fullest advantage.

New Directions

A desirable outcome of required federal program evaluation in large school systems is the movement toward the formal evaluation of all instructional programs. Through experience with Title I evaluations, it has become apparent that the usefulness of these evaluations is limited due to the lack of comparable information about all other instructional programs. This problem is especially acute in the larger school systems where it is impossible to partial out the effects of one project on a person who is being subjected to multiple influences from a variety of projects. Furthermore, interventions such as ESEA ultimately affect all other elements of the school system. These overall effects cannot be assumed to be uniformly positive. Therefore, there is a demonstrated need for continuous system-wide evaluation of all educational programs by a single research and evaluation office.

There is need for caution on the part of those who believe that research units will suddenly realize their potential for research and evaluation. Time, trained minds, and adequate supporting staff and facilities are all needed if significant and helpful service is to be given. It probably is worth repeating that in the process of solving evaluation problems, more problems will be discovered. Ultimately, however, each school district in America should be able to boast the benefits of a well defined and unified research and evaluation function--be it performed by one man or a large staff. Title I of the 1965 ESEA may be forgotten as a poverty act but long remembered as the source of systematic self-appraisal in America's schools.

Chapter 2

EVALUATION BASED ON THEORETICAL MODELS

Part I

THE COLUMBUS EVALUATION PROGRAM

I. THE EVALUATION CENTER AND ITS PRINCIPLES OF OPERATION

Introduction: A Collaborative Relationship in Evaluation

The evaluation program currently being conducted in Columbus, Ohio, is a collaborative effort between the Ohio State University's Evaluation Center and the Columbus Public Schools. The collaboration is based on joint support by the University and Columbus Public Schools, with each organization contributing both funds and personnel to the effort. The evaluation program is focused on Columbus' Title I program established under the Elementary and Secondary Education Act of 1965.

As might be expected in such a collaboration between university and public agencies, the three-year project has several general purposes, designed to meet the particular needs of both organizations involved. Among these purposes are:

> first, to provide to the Columbus Schools information to design, operate, and assess the impact of the Title I program and to make appropriate evaluation reports to the State Department of Education and the U.S. Office of Education;
>
> second, to give to the Columbus Schools a core of persons experienced and trained in evaluation, so that the school system can staff an evaluation agency within the system;
>
> third, to provide the Evaluation Center with an opportunity to

develop and test techniques, instruments, and designs for evaluation;

and fourth, to provide the Evaluation Center with a laboratory to practice in the public schools.

The general strategy for the evaluation program in Columbus has two major components. First, it includes the conduct of research and development in evaluation. Second, by using staff from OSUEC and personnel on leave from the Columbus Public Schools, it includes simultaneously provision of evaluation information for an ongoing Title I program in a school environment. Persons from OSUEC and Columbus Public School personnel gather data on the context within which projects operate, on the operation of projects from day to day, and on the impact of projects on the behavior of inner-city students. At the same time, the staff members are involved in research and development in evaluation and are being trained to take a position of leadership in evaluation upon joining or rejoining the school system.

The Role of the Evaluation Center in the Columbus Project

The Evaluation Center is an agency of the Ohio State University's School of Education dedicated to advancing the science of evaluation. As conceived by the Center, evaluation is the process of providing information which is essential to decisions made in planning, programming, and implementing activities to meet educational goals. The

Center has the following as its objectives: (1) to study planning and evaluation in education; (2) to develop models and methods for evaluation project designs, project activities, and project outcomes; (3) to develop methods and materials for implementing evaluation programs; (4) to diffuse information related to evaluation; (5) to help practitioners effectively use evaluation designs and tools; and (6) to provide instruction in evaluation. These objectives are reflected in the major activities of the Center which include research, development, training, diffusion, and limited service in evaluation.

A Conceptual Model of Evaluation

The evaluation efforts undertaken in the Columbus project were guided by a model formulated by Dr. Daniel Stufflebeam, Director of the Evaluation Center. This model is based on the concept that the purpose of evaluation is to provide information for making decisions. Pupil attainments are measured for detecting problems and specifying objectives in proposals for financial assistance; funding agencies obtain the judgments of experts about the quality of proposals before deciding to fund or reject them; project managers obtain progress reports before deciding to modify plans or procedures in process; legislative bodies require data about the effectiveness of programs before deciding to continue, discontinue, or change the level of funding for such programs. These examples illustrate that key decision makers in a change such as the Columbus program process continually need valid evaluative information. To serve such decision making requirements effectively, evaluation must be valid (homomorphic to the variables of interest), reliable (reproducible), timely (available when the decision maker needs it), credible (trusted by the decision maker and those he must serve), and accessible (available to those who have a need to know).

Since there are many kinds of decisions associated with and related to change activities, there are also different types of evaluation. The nature of decisions involved in planning and implementing change activities encompasses four general types of decisions and suggests that the needed evaluation may be conveniently divided into four generalized stages. These stages--context, input, process, and product--comprise the CIPP EVALUATION MODEL, which is depicted in Figure 1.

Context Evaluation

The first stage of the CIPP Model is context evaluation. Its major objective is to locate, delineate, and assess the importance of problems and needs in the environment where change is to occur. This is accomplished through determining, measuring, and analyzing the total system within which change is to occur. It focuses upon the setting and ranking of objectives for a system, the intended outputs,

JUNE 1967
DANIEL L. STUFFLEBEAM

THE CIPP EVALUATION MODEL

A Classification Schema of Strategies for Evaluating Educational Change

	CONTEXT EVALUATION	INPUT EVALUATION	PROCESS EVALUATION	PRODUCT EVALUATION
OBJECTIVE	To define the operating context, to identify and assess needs in the context, and to identify and delineate problems underlying the needs.	To identify and assess system capabilities, available input strategies, and designs for implementing the strategies.	To identify or predict, in process, defects in the procedural design or its implementation, and to maintain a record of procedural events and activities.	To relate outcome information to objectives and to context, input, and process information.
METHOD	To describe individually and in relevant perspectives the major subsystems of the context; to compare actual and intended inputs and outputs of the subsystems; and to analyze possible causes of discrepancies between actualities and intentions.	To describe and analyze available human and material resources, solution strategies, and procedural designs for relevance, feasibility and economy in the course of action to be taken.	To monitor the activity's potential procedural barriers and remain alert to unanticipated ones.	To define operationally and measure criteria associated with the objectives, to compare these measurements with predetermined standards or comparative bases, and to interpret the outcomes in terms of recorded input and process information.
RELATION TO DECISION-MAKING IN THE CHANGE PROCESS	Provides information for deciding upon the setting to be served, the goals associated with meeting needs, and the objectives associated with solving problems, i.e., for planning needed changes.	Provides information for selecting sources of support, solution strategies, and procedural designs, i.e., for programing change activities.	Provides information for implementing and refining the program design and procedure, i.e., for effecting process control.	Provides information for deciding to continue, terminate, modify, or refocus a change activity, and for linking the activity to other major phases of the change process, i.e., for evolving change activities.

FIGURE I

differences between intended and actual outputs, and relationships between intended and actual outputs, and relationships between actual inputs and outputs. It thus pinpoints needs and further suggests problems of system design or operation which underlie a state of need. In the change process, context evaluation provides information for deciding upon the appropriate point of entry to the change process. Decisions based upon context evaluation usually result in problem statements contained in proposals in certain critical areas.

Input Evaluation

To determine how to cope with the identified needs and problems, it is necessary to conduct an input evaluation. The objective of input evaluation is to identify possible solution strategies or coping mechanisms and to forecast their strengths and weaknesses for solving the specified problem within the relevant operating context. Input evaluation is guided by the point of entry to the change process suggested by context evaluation--whether research, development, diffusion, or adoption activities are most appropriate. The method of input evaluation is to determine the relevance of alternative courses of action to the problem of interest, their potential procedural barriers, the possibility of overcoming and the consequences of not overcoming the procedural impediments, and the benefits and costs of overcoming them. In the change process, input evaluation provides information for deciding upon program strategies and tactics. Decisions based upon input evaluation usually result in the specification of program objectives, procedure, schedule, and budget by the proposing agency and in determinations of whether and at what level to fund the proposal by the potential funding agency.

Process Evaluation

Once a planned course of action has been chosen and approved and the implementation of the plan has begun, process evaluation is needed to provide periodic feedback to project administrators and others responsible for continuous control and refinement of plans and procedures. The objective of process evaluation is to detect or predict, during implementation stages, defects in the design or its implementation. The overall strategy is to identify and monitor, on a continuous, molar, noninterventionist basis, the potential sources of failure in a project. Roger Barker has recently characterized this approach as tranducer inquiry in an ecological context.* Here, the evaluator does not exercise experimental control over the situation, nor does he manipulate it in any way. He accepts it as it is and as it evolves, and monitors the total situation as best he can by focusing his most sensitive noninterventionist data collection techniques on the most crucial aspects of the project. Such evaluation is multivariate, and not all of the important variables can be specified before a project is initiated. The process evaluator focuses his attention on theoretically important variables, but he also remains alert to any unanticipated but significant events. Under process evaluation, information is collected daily, organized systematically, analyzed periodically (e. g., weekly), and reported as often as project personnel require such information (e. g., monthly). Project decision makers are thereby provided with information needed for anticipating and overcoming procedural difficulties. A record of process information is kept for interpreting project outcomes, and in longer projects, an updating of context information--providing a dynamic baseline of needs--is a part of the process evaluation. In the change process, process evaluation thus provides information for insuring quality control in the immediate and subsequent implementation and refinement of plans.

---

*Barker, Roger G. "Explorations in Ecological Psychology," American Psychologist, January 1965, 20, 1-14.

Product Evaluation

At the end of the project or at other predetermined times, product evaluation is needed for determining the effectiveness of the project. The objective of product evaluation is to relate outcomes to objectives and to context, input, and process, i. e., to measure and interpret outcomes. The method is to define operational criteria and measure these as associated with the objectives of the activity; to compare these measurements with predetermined standards, or comparative bases; and to make rational analyses and interpretations of the outcomes in terms of the recorded context, input, and process information. In the change process, product evaluation provides information for deciding to not change, or continue, terminate, modify, or refocus a change activity, and for linking the activity to other major phases of the change process. A product evaluation might show that objectives had been satisfactorily achieved and that a developed innovation can be continued or is ready to be diffused to schools which need such an innovation (depending, of course, on the decision rules of decision makers). In another case, product, context, and process evaluation could show that while an objective had not been satisfactorily achieved, several components of the program might function differently in another instance.

From this overview of the CIPP model it can be seen that evaluation activities differ greatly in context, input, process, and project assessment phases of the total evaluation process. Both evaluation designs and evaluation instruments vary from step to step in their applicability and relative effectiveness in providing information. Many diverse types of information are needed throughout the decision-making process, and a multitude of different institutional and value constraints (such as vested interests) limit the ability of evaluation to gather that information. In addition, certain of the information may have existed prior to the evaluating effort. Such a model, therefore, demands that (1) the model be linked closely to the decision process in the program it serves, and (2) the evaluator work intimately with the decision maker(s).

Evaluation and the Decision Process in Title I Programs

The Title I program is designed to increase the education opportunities, experiences, and attainment of disadvantaged children throughout the nation by adding provisional services to them through local educational agencies. National objectives are an integral part of the program's focus as is broad control over the program's existence and support. Yet the Title I effort is specifically controlled and coordinated at the state level and is implemented at the local level. Thus, the decisions for which evaluation information must be supplied will be made at local, state, and national levels.

The loops in Figure 2 illustrate the general decisions and functions of evaluation as they may exist in programs which involve local-state-national cooperation. A set of feedback control loops delineate the national cooperation. A set of feedback control loops delineate the flow of evaluation information to levels of decision makers. The loop at the right of Figure 2 shows local school activities; the intermediate loop, state activities; and the left-hand loop, federal activities.

Block 1 portrays the local school district's program. This is the local context from which needs for educational change emerge and within which the changes to meet these needs must ultimately occur. It includes both the intended and actual inputs of the system (e. g., the learners, curriculum, staff, organization, policies, finances, physical facilities, and school-community relations), and the intended and actual outputs of the system (i. e., the cognitive, psychological, and social functioning of its students and alumni).

Proceeding to the right of Block 1, information collection is depicted by the first segment of the curved line. This collection occurs at the local school level and is essentially objective. Interpretations of the information are not relevant here, but rather the systematic collection of all information (context, input, process, and product) needed for later decisions at local, state, and federal levels. The local education agency might independently collect context information, and/or it might participate in statewide or national context evaluation, e. g., statewide testing programs, Project Talent, National Assessment. In collecting input information, the local education agency probably would seek evidence from outside the system about potential solution strategies not previously employed within the system; but information about potential solution strategies which had previously been tried in the system could also, and would probably, be collected locally. Process information about a program would be collected daily within the system. Product information usually would be collected in the local

FEEDBACK CONTROL LOOP: EVALUATION IN FEDERALLY-SUPPORTED EDUCATIONAL PROGRAMS

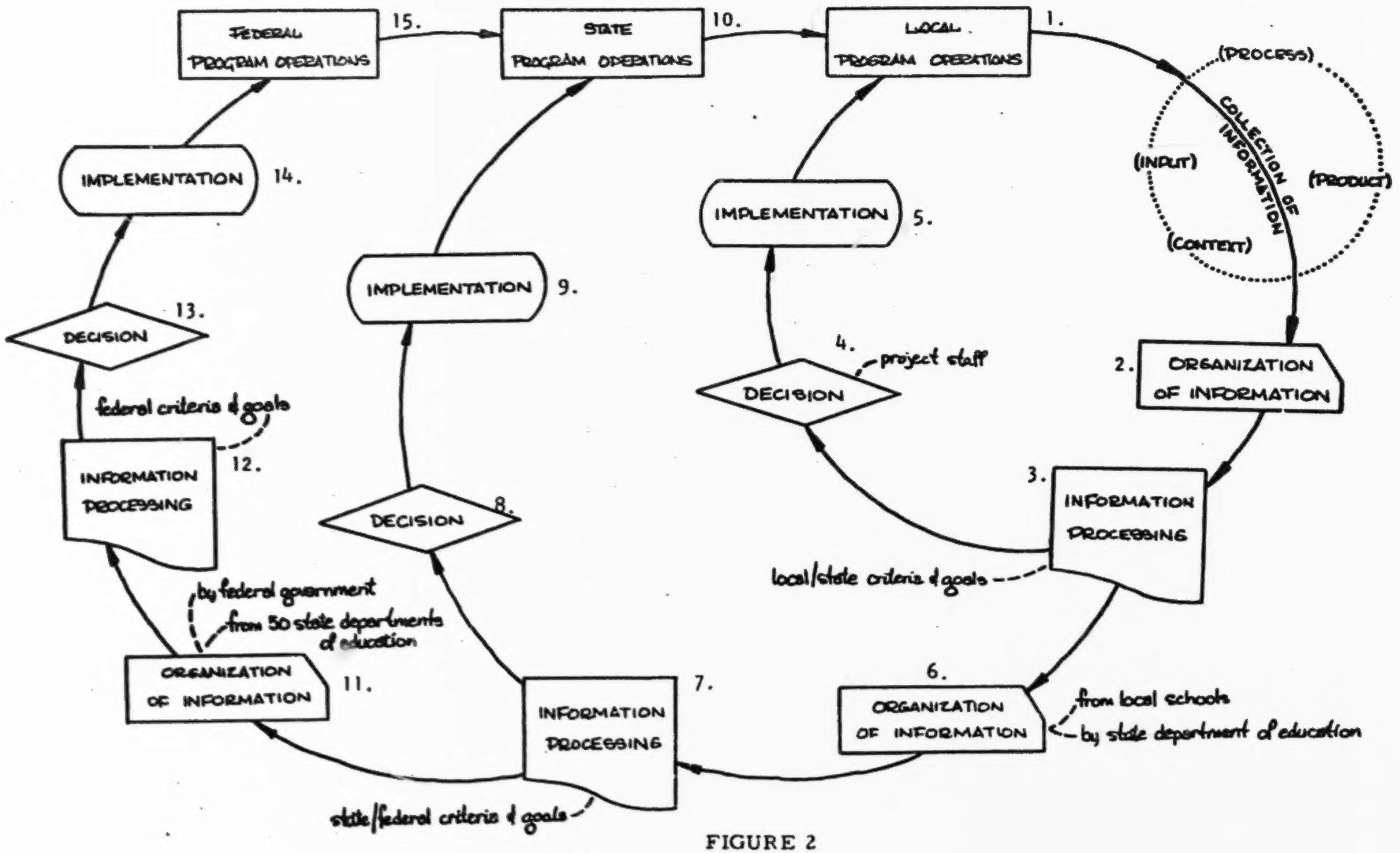

FEDERAL PROGRAM OPERATIONS
15.
STATE PROGRAM OPERATIONS
10.
LOCAL PROGRAM OPERATIONS
1.
(PROCESS)
COLLECTION OF INFORMATION
(INPUT)
(PRODUCT)
(CONTEXT)
2.
ORGANIZATION OF INFORMATION
3.
INFORMATION PROCESSING
local/state criteria & goals
4.
project staff
DECISION
IMPLEMENTATION
5.
6.
ORGANIZATION OF INFORMATION
from local schools
by state department of education
7.
INFORMATION PROCESSING
state/federal criteria & goals
DECISION
8.
IMPLEMENTATION
9.
11.
ORGANIZATION OF INFORMATION
by federal government
from 50 state departments of education
12.
INFORMATION PROCESSING
federal criteria & goals
DECISION
13.
IMPLEMENTATION
14.

FIGURE 2

setting alone at the beginning of a change activity--to provide benchmarks--and thereafter on an annual basis.

Block 2 shows the organization of information. Here information would be coded in accordance with a predetermined format, processed (e. g., keypunched), and filed regularly.

At Block 3, information collected at Block 2 would be analyzed in accordance with decision-making requirements for (1) focusing planned changes through identifying, explicating, and ranking local system needs (context reports); (2) planning strategies and tactics for effecting needed changes (input proposals); (3) directing change activities (process reports); and (4) continuing, evolving, and refocusing change activities (product reports). Information for context reports to the board of education and school administration would be analyzed infrequently (perhaps annually or as infrequently as quinquennially) to identify and assign priorities to local needs. Input information to accompany requests for internal or external support would be analyzed in the local system whenever long-range or short-range planning would be needed for seeking solutions to local system needs. Once change activities were initiated at the local level, process information would be analyzed frequently for reports of project staff to determine whether potential sources of failure were actually failing, procedures were following the project design, projects were on schedule, resources were adequate, and the procedural design or its implementation should be modified. (This information could also help to diagnose outcomes once they were determined.) Product information would be analyzed infrequently for reports to the school administration and the state education department to assess the overall effectiveness of the project in meeting its general goals and thus to provide information for reassessing the context, i.e., how have the problems and needs in the local system been modified as a result of the complete change activity?

Block 4 denotes program decisions made at the local level. Local school decision makers (including the board of education, superintendent, principal, supervisor, and/or project director) would use the evaluative reports from Block 3 to make decisions required in focusing, planning, implementing, and evolving change activities. The board of education and school administration would make decisions to focus and set priorities for change activities, and they would be aided greatly in this by access to context evaluation reports. These persons, together with a project director, would use input evaluation information for making decisions required in planning, scheduling, staffing, and funding needed programs. Following an initial context and input evaluation and the introduction of a change activity, project personnel would use process evaluation reports to make decisions required in managing a change activity. At infrequent intervals, e.g., annually, the administration would need to decide whether to continue, expand, contract, or discontinue the change activity. While product evaluation reports are important for assisting in such decisions, it should be noted that they also are invaluable aids for justifying such decisions to public and staff and for honoring those who are responsible for the successes. This probably accounts for product evaluation having reached a level of development and use which is much higher than that for context, input, or process types of evaluation.

The decisions made at Block 4 would be implemented at Block 5, and the school program would undergo frequent modification back at Block 1. This cycle is continuous. Evaluation reports would be prepared as often as decision makers in the local school needed them for program examination, planning, control, and evolution.

Returning to Block 3, product reports for the state education department would be prepared annually (or as often as required) by all public school districts in the state. At Block 6, the state education department would organize these reports into types of projects and combine the information from similar projects. This information would then be analyzed at Block 7 to determine the strengths and weaknesses of the statewide program (state product evaluation), and the state program officials would use this information to assess the statewide educational problems and needs (context evaluation) to decide at Block 8 about program emphases and state control. The state's assessment of school product reports to distinguish between effective and ineffective approaches employed by schools would constitute part of the state input evaluation. Decisions by state officials would be implemented at Block 9 and would affect the state program at Block 10. Local school projects at Block 1 would then be affected in turn.

At Block 7, annual product evaluation from the fifty states would be sent to the federal agency. This information would then be organized at Block 11 so that major program thrusts could be examined and analyzed at Block 12 on a nationwide basis and so that reports (product evaluation) could be prepared, for example, for the Associate Commissioner for Elementary and Secondary Education, the Commissioner of Education, the Secretary of Health, Education, and Welfare, the President, and the Congress. In turn, these functionaries would use the product reports to assess the problems and needs of elementary and secondary education in increasingly broader contexts (content evaluation) and to assess the effectiveness of previous plans and programs and the potential of proposed ones (input evaluation). Decisions would then be made at the federal leval at Block 13 about program emphases and funding. The implementation at Block 14 of such decisions would affect the federal program at Block 15, the state programs at Block 10, and the local school projects at Block 1.

To summarize the key points related to Figure 2: (1) information for evaluation at federal, state, and local levels will be collected largely at the local level; (2) this information will form the basis for federal, state and local decisions which will ultimately affect local operations; and (3) evaluation plans must be developed, communicated, and coordinated at federal, state, and local levels if the information schools provide is to be adequate for assisting in the decision process at each of these levels. Having presented a model for evaluation and illustrated a possible application of the proposed model in a major federally-aided program, we will next attempt to depict evaluation as conducted by the Evaluation Center within the total context of the Columbus Title I program.

The Columbus Evaluation Program: An Overview

It is hoped that many useful principles of design for evaluation programs may come from collaboration between Columbus and the Evaluation Center. The general principles upon which the evaluation was based will be outlined shortly; here, a brief overview will be presented to provide a frame of reference for viewing these principles. When viewing the Columbus evaluation program as a prototype for other or further evaluation efforts, the forces which constrain the program must be identified and understood. Several limitations on the generalizability of the Columbus program appear to exist. These constraints, both conceptual and operational, gave shape to the evaluation and are explained in the following paragraphs.

The evaluation program began formally, as mentioned earlier in this report, only after the "context" and "input" decisions concerning the projects had been made. Thus, the evaluation must remain a test of only the "process" and "product" components of the CIPP model. Nonetheless, the testing of process and product evaluation have had relevance for context and input evaluation, for many potential problems have been uncovered.

Staff members for the evaluation effort were provided both by the Columbus Public Schools and by the Evaluation Center. The school system granted leaves of absence to several of its staff members (six teachers and a school nurse), in the hope of obtaining qualified evaluators to aid in establishing an agency to conduct evaluation activities within the system. These people became "project residents," charged with the responsibility of carrying out evaluation activities for a particular Columbus project. The Evaluation Center provided personnel for the positions of director and associate director of the program. Test development specialists, data processing and statistics specialists, research assistants, and secretarial staff were funded jointly by the two agencies.

The origins of staff for the program were evident in the backgrounds and orientation of staff members. Speaking in the misleading terms of stereotypes, the staff represents people with research orientations characterized by a quantitative background and approach to problems and people with orientations to teaching and experiences with the personal concerns of students and teachers. One of the immediate changes wrought by the mixing of people of varied interests and training at the Evaluation Center was an amalgamation of their different predominant views. The result was that each staff member became sensitive to the objectives of evaluation and acquired the empathy needed to carry it out.

On the other hand, the implications of staff differences were not immediately evident. For example, only through painful experience did it become obvious that any future program to train evaluators must be designed only after assessing the capabilities and viewpoints of trainees and trainers participating in the training. A context and input evaluation of the evaluation training program might have pointed out the importance of this problem.

The data generated by the evaluation effort have remained in the ownership of the Columbus Public Schools. Such a policy is in accord with the recognition that evaluation ought to be oriented to decision makers in the program. (This is explicitly recognized in the conceptual relationship between evaluation and the decision process in the Title I program.)

General Evaluation Design Principles

Principles serving as a foundation for the Columbus evaluation program are concerned with the design of process and product evaluation activities, the construction and administration of instruments, and the manner in which obtained evaluation information is fed back to the project staff.

1. Orientation meetings are held at the beginning of projects and periodically during the life of a project. The purpose of these meetings is to inform the participants in a project of its objectives, its procedures, and its evaluation. At orientation meetings, a special point is made of discussing the purpose of evaluation, with emphasis put on the need to improve the project by continually evaluating it. By such means, the evaluation staff attempts to involve others in the process of evaluation and to diminish its threat. Such meetings also give personnel in the school system a chance to contribute to the design of a project and its evaluation--or at the least to suggest potential barriers to the success of a project so that these barriers may be monitored through process evaluation.

2. Initial design of the evaluation effort is a collaborative activity. Objectives for each of the projects in the Title I program are set and written by the appropriate project staff from Columbus Public Schools. During and after the period of setting objectives, the director of the Title I project under discussion has the responsibility for detailed project objectives. The Evaluation Center staff has the responsibility to review, comment on, and act on these objectives, striving to clarify objectives and couch them in terms of observable behavior. At the

same time that objectives for a project are determined, criteria for measuring objective attainment are set for use in product evaluation. Thus, an overall design covering each project for a full school year is written by the Evaluation Center. Finally instruments to fulfill the demands of the design are chosen or constructed by staff members of the Center with periodic consultation being provided by the University's Test Development Center.

3. Process evaluation activities are given focus and defined prior to beginning Title I project activities in the schools. The Columbus Title I project staff and project residents from the Evaluation Center jointly suggested possible barriers to the success of the project. These barriers then served as the focus of process evaluation, as the project was continuously monitored. If information was generated on other barriers to project activities, the process evaluation was then refocused, instruments were revised or new ones constructed, and the evaluation proceeded.

4. Process evaluators are noninterventionists, but process evaluation is intervening. A guiding principle in the evaluation was that evaluators themselves should not intervene in the operation of a project. To meet this objective, much of the needed process information was collected from logs, records which were needed for purposes of bookkeeping, and unstructured and conversational interviews. However, the assumption upon which process evaluation itself is

predicated is that evaluators serve as decision makers, and, to the extent that process evaluation causes shifts in project operations, it does indeed intervene. Hence, a guiding principle of process evaluation is that it intervenes to guide the project, although this intervention is at the discretion of project decision makers.

Unfortunately, it is especially difficult to determine how, or if, process evaluation and evaluators intervene in the project. Evaluators who are involved in their jobs, interested in providing process information which reflects the true operation of a project, and concerned that data are used to make decisions and verify their appropriateness--such evaluators interject unknown influence into the project's operation. Often process evaluation _is_ obtrusive, in spite of efforts to prevent intrusion. While such (quickly-designed) devices as multi-project interview schedules, the use of non-overlapping samples of respondents, and the tight scheduling of process instruments mitigated the impact of the evaluation on school personnel, nonetheless the evaluation did consume _their_ time and effort as well as that of the evaluator.

5. _One of the guiding principles of this evaluation has been the guarantee of anonymity to persons when they respond to instruments._ An assumption in conducting the evaluation has been that participants in _any_ project are _competent_ in their roles. Thus, if problems occur in day-to-day operations, or if the hoped-for product is not produced, some component of the project itself is assumed to be at fault. To

make this point clear to respondents, it has often been stated (so as to diminish the hesitancy of persons to respond) that people are not being evaluated.

The guarantee of anonymity has, in several instances, been a point of contention between evaluation and project staffs. The probability of contention appears especially great when individuals, or characteristics of a project which bear a person's vested interest, are identified as barriers to success. At times, project staff have requested information on the identity of persons who responded to instruments or provided a particular piece of information. However, denial of such requests and reaffirmation of the principle of anonymity has proven to be essential to continually obtaining adequate information for the Title I program.

6. Data are obtained by means of sampling. This approach has been taken for the sake of insuring some degree of reliability and validity and has been employed in both process and product evaluation activities. (In addition, sampling served as a strategy to keep percentages of response to an instrument high, by not overburdening one respondent. Often several samples of the same class of respondents were used for different instrument administrations.) Principles of random and stratified random sampling were employed when possible in the selection of schools and persons to whom instruments would be administered. The sampling approach and the rather large number of

instruments administered to project participants have made it necessary to schedule instruments carefully for administration in the schools. In addition, it was found expedient to combine instruments from different projects into a single jointly-used instrument, when the instruments attempted to measure similar objectives present in several projects.

7. Both commercially-available and specially-constructed instruments are used in the evaluation. When possible, instruments normally administered as a part of the Columbus Public Schools testing program are also used. Among the specially-constructed instruments for the project are depth-study interview schedules for individual children; questionnaires and interview schedules for teachers, administrators, and special personnel assigned to projects; Q-sort instruments; attitude surveys, rating forms for facilitating factors and deterrents to project success; and various logs and record forms. Project residents construct or assist in the construction of these instruments as part of their training.

8. Feedback sessions are held monthly during the life of a project to inform the project director (and his staff) of the operations of the project. Such sessions are attended by project residents from the Evaluation Center and by project staff. Process information is reported verbally and by other media (films and charts). Project staff members are then given an opportunity to ask questions about findings or

participate in refocusing the process evaluation. In addition to these sessions, meetings are held between Columbus' Assistant Superintendent for Special Services and the directors of the Evaluation Center and the Columbus evaluation program within it. The program (all projects), the operation of evaluation activities, and the planning of future activities can be discussed at these meetings, while at the same time separate channels of communication between the University and the school system remain open (through the upper and lower levels of the system).

The flow of evaluation information in the school system's channels of communication at times presented a problem. Feedback, while designed for all project participants, often reached only the upper levels of project staff. Evaluation information was provided to top-level decision makers and had to flow from the top of the organization downward. When passage of information was impeded, or when only certain information was passed as desired, lower-level personnel felt they were obtaining little benefit from the time and effort they contributed as respondents in the evaluation. It would appear that all evaluation information ought to be channeled to all project participants in some form or other and often enough to retain their sense of involvement in the project and commitment to the evaluation.

Evaluation, after all, can only be timely when access to information is guaranteed to those project staff members who have a "need to know." It is apparent that the manner in which information is to be disseminated must be determined in the early stages of the evaluation, for dissemination itself is a strategy for improving activities by influencing decision makers.

The Role of the Evaluator

The role of the evaluator is a technical one demanding a high level of skill in the use of techniques for collecting and organizing, analyzing and reporting, information concerning the project he is evaluating. Also, the evaluator must be cognizant of the information requirements of those persons who must make decisions in and about Title I programs and use these requirements in designing and implementing evaluation activities.

However, it appears that the evaluator operating in a school system may have a more difficult role to fulfill than is apparent from the previous description. The evaluator must provide, for decision-making by the project staff, information which is credible, useful, valid, timely, and reliable. The evaluator, therefore, must spend much of his time working with (or among) project staff. However, it is apparent that the project staff is only one group--a part of the local education agency--of several groups which must make decisions concerning Title I activities. It is quite possible that conflicts may arise over the nature of information needed. The evaluator may desire to collect information for other levels of decision makers involved in the Title I program, e.g. the State education department, while the project staff may dispute the necessity or the timing of collecting such information. Questions about the credibility and objectivity of evaluators almost invariably arise from the evaluator's clients in many quarters whenever the agency being evaluated imposes limitations on the data to be collected. In such a circumstance, resolution of conflict is possible only through engendering understanding about what all relevant decision makers need.

A second potential problem inherent in the evaluator's role is the requirement that the evaluator make no decisions himself for the project. The evaluator has the ability to subtly shape the path and progress of Title I activities by his use of information. Thus, the evaluator must affirm often after much searching of his conscience that he will take a neutral position with respect to the direction of the project, except in technical matters of evaluation. To the extent that evaluators deviate from this neutrality, unknown intervention in project activities has begun.

Research Assumptions, Efforts, and Special Studies

Research into the theory, methodology, and usefulness of evaluation is a key objective of the Evaluation Center's overall strategy for evaluation in the Columbus Public Schools. However, while research into the area of evaluation is valued by the Columbus Public School system administration, it is certainly true that the administration has research interests of its own. Thus, while the schools may value research, they must balance all demands for research in the school setting, as well as concerning themselves with how true an image of the schools such research presents. In short, research in an ongoing evaluation effort appears to be a particularly sensitive endeavor: this is especially true when inquiry is not so much concerned with possibilities (the intent of the experimental investigator, who asks "What would happen if...?") as with actualities(the intent of the aexperimental investigator, who asks "What does happen in the real world?"). Most of the efforts of the Evaluation Center--although it is not concerned with experimental research on the school system's problems it uncovers--have produced context and process data. These data by their very nature are threatening, for they portray problem areas, shortcomings in previous programs, barriers to progress, and deterrents to success. Explicit research efforts of the Evaluation Center (discussed shortly) often produced data concerning the way evaluation information was utilized and the effect it had on the programs of concern. Such research has the power to threaten and at times may have touched sensitive points in the school system.

The combination of threat which can come from research efforts, and the policy of vesting ownership of data with the Columbus Public

Schools, may have caused a general uneasiness about research on evaluation in conjunction with the evaluation program. If a general principle can be formulated, it is that research efforts must proceed tactfully, and proceed only after their rationale and needs are firmly established. The alternative to this cautious and rather slow course of action is to bargain, before the evaluation, for the type and amount of research to be conducted concurrently with evaluation services.

A number of research projects have been produced as an integral part of the Columbus evaluation program. Most of these projects have been concerned with the state-of-the-science of evaluation, its potential usefulness in the school, and some of the ways in which evaluation may be more efficiently conducted. Among those efforts, proceeding or completed, are studies of the roles and training of evaluators, the use of information management and management information systems in education, the relation of evaluation to decision making in planned educational change, the effectiveness of item sampling techniques as a method of estimating test norms, and a study on the use and scoring of the Color Pyramid Test. A bibliography on evaluation has been compiled with sources ranging from classical research design to contemporary views of evaluation. In addition, a critical survey of the literature in evaluation has been completed to serve as a base of theory for conceptualizing further evaluation efforts.

Ruth Odgers of the Evaluation Center conducted a brief study on "The Selection, Training, and Functions of ESEA Title I Project Evaluators" (available in mimeographed form from the Ohio State University Evaluation Center). The objectives of the study were to determine: (1) the functions of a project resident or member of an educational evaluation staff, (2) the type and amount of experience considered essential for performing such functions, (3) the training believed to be necessary, and (4) the personal traits believed necessary for evaluators. An open-ended questionnaire was administered to a selected group of key school administrators from the State of Ohio and to members of the Evaluation Center staff.

There appeared to be three general agreed-upon functions of the evaluator: (1) planning educational evaluation; (2) implementing educational evaluation plans; and (3) disseminating evaluation findings. As a professional educator, the evaluator was expected to assist school staff in the process of conceptualization and determination of educational goals, plan programs for their attainment, and help school staff assess progress at given points in the program. Unique attributes of the evaluator's role were (1) the designing of evaluation programs; (2) constructing and/or devising instruments for data collection; (3) analyzing and organizing data; (4) reporting findings; and (5) giving help in using results for the purpose of implementing, terminating, or modifying projects so that goals could be effectively attained. The type of experience and training believed to be necessary for carrying out these functions included a Bachelor's degree in education; graduate casework in educational research, and preferably a Master's degree; and certification in teaching. Personal traits believed important, in addition to those customarily-mentioned ones such as a liking for people, a quick intellect, stability, cooperativeness, and concern for accuracy and objectivity, were flexibility and willingness to compromise and take risks; a tolerance for ambiguity; a high frustration level; tact; and the capacity to accept and utilize criticism.

Because of the orientation of the Evaluation Center's efforts to the needs of decision makers, it appeared necessary to construct a taxonomy of the decisions which project directors would make. Jack Ott, now director of the Ohio State University Test Development Center, constructed such a taxonomy and tested it by observing the behavior of project directors in Columbus during the course of the year. This work, serving as the basis for his dissertation, had two major conceptual components. The first component included the premises, anatomy, and typology of educational change. The second component included a taxonomy of the decision process which was observed to be part of planned educational change. From these two conceptual sources, a classification of decisions that project directors made was derived. Three means were used to gather empirical data which would both support and interact with the construction of the taxonomy of decisions. First, a log was constructed and used by project residents to note the decisions made by project personnel. Second, monthly feedback meetings were observed and the decisions made during the course of meetings noted. Third, substantive reports of evaluation efforts were analyzed to find decision makers' need for information concerning their premises. To test the usefulness of this decision taxonomy, the questions, information needs, and decisions which were made by project personnel were summarized by projects. (The outcome of this study will soon be available from the Evaluation Center; interim findings can be obtained from Jack Ott, Director, Test Development Center, The Evaluation Center, The Ohio State University.)

More effective means of conducting evaluation activities were explored in an experimental study undertaken to determine and compare the validity of item sampling with examinee sampling for estimating test norms. Examinee sampling was defined as the case in which norm statistics were computed from data obtained by administering all items in a test to a random sample of examinees; item sampling, on the other hand was defined as the case in which one set of norm statistics is derived from several non-overlapping sets of test data, each of which is based on random samples of items and examinees.

The study was designed to answer the following questions:

1. Would item sampling, where each of n examinees is administered only a proportion of the items from a single test, yield as good or better a norm estimate as examinee sampling, where all items are adminstered to some proportion of the examinees?

2. Was the effectiveness of item sampling a function of item sample size?
3. Is item sampling as valid for establishing norms for "advantaged" as for "disadvantaged" children?
4. Within limits, does a large number of examinees taking a small number of items each, or a small number of examinees answering many items each, give a more accurate norm estimate (keeping the product of examinees x items taken constant)? Two unique features of this study are that it employs item samples as small as 6 percent of the entire test, and that it involves elementary school students.

The basic study was replicated in two public school populations: 1665 subjects from 77 fourth-grade classrooms were chosen from schools in "disadvantaged" neighborhoods, and the same number from schools in "advantaged" neighborhoods. The instrument employed was the Word Knowledge subtest of the Metropolitan Elementary Reading Tests. Means, standard deviations, and item statistics were computed for each item in the item and examinee sample.

The findings regarding the first study questions indicated the item sampling approach performed about as well as, if not better than, the examinee sampling approach. These findings support item sampling as acceptable, and probably preferable, as an alternative to examinee sampling for estimating test norms. Findings regarding the second study question are inconclusive: certain findings suggested that the accuracy of the estimated means increases as the size of the item samples increase; however, other findings suggested that exactly the opposite was the case. It also appeared that a combination of small item samples taken by a large sample of examinees produced the best norm estimates. These findings will have to be examined carefully, however, before conclusions can be drawn. The findings are again inconclusive for the third study question; they suggest that sampling overall provided a more accurate estimate for the mean and variance for the disadvantaged study population than for the advantaged one. However, the differential effectiveness of item sample norm estimates for advantaged versus disadvantaged study populations was slight; thus, it would appear that item sampling as a technique is at least as effective as examinee sampling for advantaged and disadvantaged study populations. Regarding the fourth question, it appears that a combination of a large number of examinees taking only a few items each produces more accurate norm estimates than a combination of a small number of examinees taking many items each. (This research is more extensively reported in "An Experimental Comparison of Item Sampling and Examinee Sampling for Estimating Test Norms," by Thomas R. Owens and Daniel L. Stufflebeam, mimeographed by the Ohio State University Evaluation Center.)

A feasibility study of the use of information management and management information systems was conducted by several members of the Evaluation Center staff. The study, much like the one concerned with the relation of evaluation to decision making and planned educational change, was concerned with the type of information needed for project decision making. In addition, it was also concerned with information needed for decisions at the local school system and state department of education levels of administration. Its basic objective was to determine the feasibility of anticipating these information needs prior to the necessity for making decisions.

Questions which were asked by decision makers were examined both formally and informally. Decision and question content and frequency were assessed to determine what, and how often, people felt they "needed to know." Two general strategies for providing information emerged from the study: (1) the strategy of anticipating, on a stand-by basis, the demands which decision makers would have for information; and (2) the strategy of providing a general base of information to be used in decision making.

The major finding of the study thus far, in terms of feasibility, is that not enough is known concerning the questions which decision makers ask to anticipate the information needs which decision makers will have. Evaluators involved in the study now believe they know enough to ask questions concerning decisions, but the decisions themselves and their frequency of occurrence are seldom known. It appears, and reasonably so, that decision makers do not ask questions until they feel a need for information; however, it is also true that given the present state of the art of evaluation, a certain amount of "lead time" is involved in providing information for decision makers.

In addition, it may be part of the evaluator's role (if he so chooses) to attempt to help the decision maker in anticipating the decisions he must make in the future.

It is necessary to balance the requirement that information be provided quickly with the requirement that in order to provide information of high certainty a certain amount of time must be consumed in gathering, analyzing, and reporting information. Given the necessity of trading off between demands for time and accuracy, one strategy which appears viable at the present time is to utilize secondary data sources rather than primary ones in gathering evaluation information. (Some of the thoughts which motivated the conception of this study are contained in "Evaluation of Planned Educational Change at the Local Education Agency Level," written by Howard O. Merriman, and published as Occasional Paper 67-106 of the Division for the Study of Education, School of Education, The Ohio State University.)

Two research efforts have been focused on literature which currently exists in the field of evaluation. The first work, a bibliography on evaluation materials available in the library of the Evaluation Center, is concerned with gathering references in education research, organization theory, conceptual work in evaluation, innovation, change, decision making, methodological works on the design of evaluations, the gathering and analyzing of information, and the reporting of information to decision makers. This bibliography is constantly updated, and supplements are issued to keep evaluators abreast of emerging work in the field. The second work is a paper prepared on "A Selective Review of Literature in Evaluation," by Mary C. Burger (mimeographed by the Evaluation Center). The review scans evaluation literature in order to identify trends, highlights, and emerging views of evaluation. The aim of the review is to provide a conceptual base upon which Evaluation Center staff members can build further theories of evaluation.

One of the major, but unstructured, thrusts of research in the Evaluation Center during the course of this project has been conceptual research into the context, inputs, process, and product of evaluation as it relates to the field of education. The CIPP model has been expanded, explicated, analyzed, and dismantled many times, formally and informally. While such work cannot be reported as a particular paper or set of guidelines, the conceptualization and re-conceptualization occurring at the Evaluation Center is manifested in the efforts and papers reported in this discussion. In addition, the CIPP model has been expanded in its application, being applied to the Bureau of Elementary and Secondary Education's programs in the U. S. Office of Education and to the Title III program under ESEA of 1965.

An area of special interest is the area of instrument development. This area has received a considerable amount of the creative and administrative energies expended during the past year by the Evaluation

Center. In addition to the development of such instruments as questionnaires, interview schedules, attitude scales, and reporting and accounting forms, several quite unique instruments have been developed for use in the Columbus Program. The instruments include a Q-sort for use in determining the perceptions of the role and function of personnel on various projects from the viewpoint of other personnel; an interview and questionnaire schedule for simultaneous questioning of a respondent concerning a number of projects with similar components, participants, or objectives; a set of objective scoring techniques for the Munsterberg Incomplete Stories test, which can be used by staff members unskilled in the use of this instrument; a facilitant-restraint rating scale, which allows personnel in projects to indicate the degree to which certain factors of the project impede or expedite the operation of that project; and preliminary conceptual work on such instruments as a cultural awareness test and an observation scale for use in operating classroom situations.

In summary, instrument development has been a major thrust of the Evaluation Center's efforts, especially since few instruments are available which can gather information required by designs for process evaluation. (Additional information and copies of instrumentation used in the Columbus project may be obtained from the Evaluation Center.)

II. EVALUATION IN THE COLUMBUS PUBLIC SCHOOL TITLE I PROGRAM

The Columbus Title I Projects and Program: An Overview

A brief look at the context of evaluation in Columbus--the Columbus Public Schools' Title I projects and program--may be helpful in assessing the Evaluation Center's activities. Several characteristics of these Title I projects have come to shape the nature of evaluation activities. First, while all eight of Columbus' projects focus on meeting particular acute needs of the disadvantaged child, they do so without full integration of their efforts. The Columbus program is not a single project with many component parts but one objective--a child with *n* qualities. It is rather an array of projects each focused on specific problems common to disadvantaged children in certain groups (though certain of these programs complement one another). Columbus has chosen to work intensively on particular problems of different groups, rather than provide what *might* be a less effective base of uniform services for all children in the target population. The implication of this choice for evaluation activities is that a "total system" view must be taken of individual projects themselves rather than viewing all projects together. This also has had an operational impact in the Evaluation Center--a separate evaluator (or team) works with each project and its director.

A second characteristic of the project which shaped evaluation activities is that most of the context and input decisions had already been made by program designers prior to the involvement of OSUEC. The problems were identified, inputs assessed, and the program designed before the evaluation effort was launched. However, the evaluation staff was involved (often intimately but to varying degrees) in setting operational objectives and the criteria for measuring their attainment, as well as in formulating process evaluation concepts with project staff.

The General Context of the Projects and Program

The Columbus Title I effort is made up of eight projects operating within the environment of a large metropolitan area. To place the program and its evaluation in perspective, Columbus is a city of almost 600,000 people in a county of over one million residents. The Columbus Public School system reflects this size, as it encompasses approximately 105,000 students and 112 square miles. (For purposes of comparison, Columbus' system is the second largest in the state, after Cleveland, and just within the top 20 largest school systems in the country.)

The city is plagued with many of the problems which were an impetus for passage of the ESEA of 1965. Many of its problems may be described by stereotyped urban-suburban disparities. For example, there are a large number of non-whites and low-income groups in the central city and a smaller number in the urban fringe and suburbs around the city. However, an atypical aspect of Columbus' problems concerns aspiration. Columbus, the capitol of the state, has an economy based largely on government and service industry: a large proportion of its labor force is employed by state, county, and local governments, by the Ohio State University and other universities, by financial institutions, and by other industries which provide services for the region. For these reasons (among others), members of minority groups often have challenging and satisfying jobs. Many of them are also highly educated and are vocal in their concern over the plight of those in need. Thus, high levels of aspiration are generated among minority group members and "disadvantaged" persons, but frustrated by life in Columbus' economy.

The concentrations of "disadvantaged" families and children in the central city and the aspirations of these residents have placed pressure on the public school system to provide services and facilities to enable these disadvantaged people to obtain the skills they need to become participating members of society.

The Columbus Public School system has structured its organization (as well as its Title I projects) in an attempt to meet the educational needs of disadvantaged and inner-city children. A Division of Special Services has been established and is headed by Dr. Joseph L. Davis, Assistant Superintendent for Special Services. It is the purpose of this Division to plan for and guide educational activities which are designed for inner-city children. The Assistant Superintendent for Special Services, then, has the responsibility for operating and evaluating the Title I program in the schools. In addition to this organizational feature, the system has recognized the existence of several degrees of need for special services and has assigned different priorities to these services (this integral part of the Title I program will be discussed later). The school system has no division for research and development; however, it draws freely on the services of Ohio State and other nearby universities for such advice and assistance as it may need.

During the 1965-1966 school year, the Columbus Title I projects involved almost 8000 public and over 600 private school students clustered in 34 public and 5 private schools across the city. To serve the needs of these students, over $2, 150, 000 of Federal assistance was made available during the 1965-1966 and 1966-1967 school years in the form of Title I monies. As a first attempt to create the greatest benefit from these monies, the needs of schools were answered. For this assessment, categories of need--Priority I, II, III and IV--were established. Each of these categories represented schools in neighborhoods which exhibit respectively less and less need as determined by the measures presented in Title I guidelines. Relevant records on pupil locations, achievement, and problems were chosen and used by the school administration in this assessment. Following this assessment, overall objectives were set, projects were designed, and resources were allocated in consideration of the type and degree of need exhibited. It is important to note that the evaluation program, as it is discussed further in this report, did not come into being until after these decisions about need were made.

The Title I Projects and Evaluation Instrumentation

Title I projects operating in the Columbus program include After-School Study Centers Projects, Basic Mathematics Improvement and Reading Improvement Projects, Elementary Counseling Project, Enrichment Unit (language arts) Project, Health Service Centers Projects, Pre-Kindergarten Project and Regional Service Centers Project. The project activities and the staff who perform them are described in the following sections. Each of these projects includes activities for children and for teachers. While the activities are oriented to the known needs of disadvantaged children, their needs are not fully known--thus the projects are evolving. Since many teachers were not fully aware of the breadth and depth of these children's needs, the Columbus projects each include an in-service training component designed to increase the teacher's understanding of the child and his environment. The in-service training program may be, and the efforts of the Evaluation Center must be, recognized as sources of feedback to the system and as attempts to tailor education more and more closely to the needs of these children.

The Pre-Kindergarten Project of the Columbus Public Schools is designed to provide an organized program, medical services, and social services for socially disadvantaged children who are about to enter kindergarten (or first grade the next school year). It is focused on children whose home environments suggest great need for pre-kindergarten experience. These children are usually four years of age. Approximately 85 percent of the children come from homes with incomes of less than $3000 a year.

The program is designed to enable children to develop an adequate set of initial skills and the disposition toward intellectual tasks required for educational progress. Strategic objectives include the development of perceptual skills, linguistic skills, and a mental set which might be called "learning to learn"--including motivating the child to find pleasure in learning; developing his ability to give attention to others, to pursue purposive action, and to delay gratification of his desires and work for more distant goals; and enabling him to view adults as sources of ideas as well as sources of approval and reward. In addition, the program involves parents by welcoming them to classes for observation and having teachers visit with them and discuss the progress of their children and the role of the home in this progress. Also, children served by this program are given a health examination (including physical, dental, and visual components). Immunizations are available through school health services. Referrals are made to consultants or health facilities in the area, and these referrals are followed up by social workers, aides, and school nurses.

The program operates in nearly twenty centers in inner-city target areas during the school year (mornings and afternoons). The staff assigned to each center consists of a teaching and social service team composed of certificated kindergarten-primary teachers, lay helpers and volunteers (often residents in the target area), and two social workers (usually shared among several centers). Also, health service personnel from the Health Center Project (described later) provide appropriate services.

The instruments for the Pre-Kindergarten Project were all developed at the Evaluation Center for use in the project. One of these is a child study record, used by the teacher to check each child's health, muscular coordination, emotional stability, social adjustment, general intellectual capacities, art skills, dramatic play, music and rhythms and language skills, and a checklist for the treatment of problems which the child manifests. (The child study record is so designed that its information can be transformed into an optical scanner input sheet and be processed by computer.) Also included are an accounting sheet for enrollment and the number of students on waiting lists; a materials rating scale for those items supplied for the pre-kindergarten student; interview schedules for parents, teachers, and principals involved in the project; an accounting sheet for enrollment and attendance records; and a supplemental log sheet for successes, problems, and recommendations which the teacher feels are significant.

The Enrichment Unit Project is the largest of Columbus' Title I projects, serving over forty public and private schools in the inner city. There are two major components in the program: a language arts program for primary-level pupils; and an administrative structure for primary teachers, designed to allow these teachers to make more productive use of their time.

The language arts program is constantly evolving to meet new needs of disadvantaged students and to make the most effective use of new discoveries concerning these children. Process evaluation information, the observations of participating teachers, study teams concerned with reading, and consultants from outside the public school system are all involved in this evolution. Specialized materials have been introduced which are oriented to the life styles of students, and experiences designed to broaden the cultural awareness of disadvantaged children have been provided.

The administrative structure of the program, the "enrichment unit" is based on providing an enrichment teacher for every three or four primary teachers and on creating a system of flexible scheduling so that enrichment and regular teachers may work together in various ways. The enrichment teacher, highly skilled in the teaching of language arts, is used to support the language arts instruction of the regular teacher. Other major benefits are freeing the regular teacher for more frequent and productive visits to students' homes and allowing the regular teacher to grow professionally by having more time for study and lesson preparation.

Three schedules involving the enrichment and regular teachers have been established. One permits the enrichment teacher to assist a regular teacher with language arts instruction, spending the same proportion of each school day in each classroom. A second schedule permits the regular teacher to relinquish her duties to the enrichment teacher for one day each week; the enrichment teacher would then have responsibility for the entire instructional program in the classroom. A third schedule relieves the enrichment teacher of her classroom duties.

Each of these schedules results in different benefits for the teachers and students involved. The first allows for general flexibility in the classroom by teachers to vary the student-teacher load, to break the students into groups for discussion, and to take students on trips to engage in outside activities. The second permits the regular teacher to engage in home visits,to participate in professional growth activities, and to develop lessons and language arts materials needed for her students. The third allows the enrichment teacher to assist in the development of a language arts program or a program of parent education or to participate in professional growth activities.

All the instruments, except one, for evaluation of the Enrichment Unit Project were developed at the Evaluation Center. The (non-standardized) instruments developed include (1) a classroom teacher interview, to sample the teacher's opinions about the projects's attainment of its objectives, elicit recommendations for improvement of the project, and determine barriers to the success of the project; (2) an enrichment unit teacher interview schedule, to obtain the same information as above from the enrichment unit teacher; (3) a principal interview schedule, eliciting information on communication and co-operation in the project, the progress of the home-school program, professional growth and language programs (as well as soliciting recommendations for the improvement of the project and information concerning the barriers to its success); (4) a combined principal interview with pre-mailed questionnaire; (5) a general recommendations log for gathering classroom and teacher opinion concerning effective and ineffective practices in the enrichment unit; (6) a Q-sort developed to examine both the ideal and actual role of the enrichment teacher as perceived by enrichment teachers, classroom teachers, and principals; (7) several materials rating scales to gather information about the organization, distribution, and use of selected materials designed for the project; (8) a language program recommendations log for gathering information concerning the language program component of the

Enrichment Unit project; (9) a professional growth questionnaire for determining the effect of and reactions of teachers to the professional growth meetings; (10) a home-school relations activities inventory and a home-school relationship recommendations log; and (11) a parent interview to gather information on parental understanding of the project, parental perception of the effectiveness of the project, and parents' opinions concerning home visitations. In addition, the Cooperative Primary Tests of Language Achievement (written by Educational Testing Service) were used in a pre-test-post-test design to assess pupil achievement in the language program.

The Regional Services Center Project is designed to improve the efficiency and effectiveness of providing educational opportunities for disadvantaged students. Service centers have been established for regions encompassing five elementary schools where substantial numbers of disadvantaged students are enrolled. These regional service centers are staffed with subject-matter specialists to provide coordinated curricular services in elementary-school science, foreign languages, art, music, and physical education. In addition, the centers are equipped to provide resource materials for these five curricular areas. Activities specific to these curriculum areas are designed with the unique needs of the disadvantaged student in mind. Staff members are chosen for their quick competence in subject-matter areas and for their ability to adapt activities quickly and make them understandable.

Approximately 30 schools in the inner city are served by regional service centers. Each center is staffed by 20 teachers: five are resource teachers and three represent each of the five curricular areas. In addition, elementary teachers from the Mathematics Improvement Project and elementary counselors from the Elementary Counseling Project are housed at the centers. The staff members rotate out of the centers and through schools in the "region," usually on a weekly basis.

Each of the evaluation instruments administered to participants in the Regional Service Centers Project was constructed by members of the Evaluation Center staff. Included among these instruments were the following: (1) *a classroom perception scale,* to collect information on the perceptions of teachers concerning the functioning and impact of resource teachers (from regional service centers) in the classroom; (2) *a data collection sheet* focusing on the hours spent and number of classrooms seen by each resource teacher; (3) *questionnaires for general resource services, science services, visual arts improvement services, foreign language services, music resource services,* and *physical education services*, to gather information on the role of the resource teacher, as well as information from the resource teacher concerning training, coordination of activities, communication, services provided, materials used, and recommendations which seem significant; and (4) *a principal interview* focusing on the interaction between the regional services center and the principal's school and the utilization and effectiveness of resource teachers provided.

The Elementary Counseling Project is designed to provide counseling services for disadvantaged students in the target area of Columbus' inner city, represented by more than 28 public and private schools. The project is also designed to provide counselors who are certified and qualified for elementary school work in order to continue and expand the project's activities.

Counseling activities, focused on grades 1 to 6, aim to assist each student to better understand himself, see his relations to the world around him, and understand his responsibilities in interpersonal relations and in school. In addition, counseling serves as a means of identifying problems for correction and attempts to improve the school's communication with parents.

Seven counselors comprise the staff of this project. Persons with a successful elementary or counseling experience are selected. Following an intensive orientation to their jobs, they are assigned to rotate through four or five schools. They provide approximately one day's service per week to each school. The counselor's final day each week is divided between working with other staff and pursuing university coursework toward certification. Similar services are provided to private schools, with students referred to the counselor by staff members.

Both standardized and non-standardized instruments were used for evaluation activities in the Elementary Counseling Project. The standardized instruments employed were the Revised Munsterburg Incomplete Stories, a Social Acceptance Scale, and the Color Pyramid Test. (In addition, a set of objective scoring guidelines were developed for the Munsterberg Stories, so that they could be scored, from tape recordings of the administration and subject's responses, by personnel without previous experience in the use of such instruments.)

A number of non-standardized instruments were developed for the project by the staff of the Evaluation Center. Among the instruments developed were (1) accounting sheets for individual counselors (to obtain information concerning professional activities which had engaged staff members), for project directors (investigating plans for expansion and guidelines used for the project), and for parent contacts made by counselors for differing purposes; (2) a counselor interview guide, to guide interviews (with counselors) which focused on the operation of the project and the counselor's perception of it; (3) a facilitant-restraint scale to determine factors of the project which made project operation easy or impeded it; (4) a Q-sort to determine the ideal and actual roles of the counselor as perceived by himself, teachers, and principals (together with a sorting sheet to allow the Q-sort to be self-administered privately and mailed in for analysis); (5) a principal and teacher questionnaire; (6) a professional growth questionnaire to gather data concerning the information gained from professional growth meetings which proved useful for counselors in their daily work; and (7) an inventory of guidance awareness to obtain the perceptions of teachers concerning the elementary counseling project.

The Mathematics Improvement Project is designed to provide a structured mathematics program for pupils in the inner city who are not achieving at a level commensurate with their ability. Highest priority is given to pupils who are farthest below their grade. Both elementary and secondary school pupils are included in the target population of this project.

The basic objectives of the program are to increase each child's fundamental computational skills, help him develop a vocabulary which will aid in understanding and expressing mathematical ideas in his daily life, develop his ability to think logically and to solve problems in a quantitative situation, and motivate the student to expend greater effort to perform consistent with his ability.

The project includes seven mathematics improvement teachers assigned to junior high schools in the inner-city area working with grades 7, 8, and 9 and five elementary mathematics teachers assigned to regional service centers working with grades 4, 5, and 6. Students are identified by their scores on mathematics tests normally used in the school system and are subsequently referred to mathematics improvement teachers. Schedules are arranged (for example) so that junior high teachers work with a different group of underachievers each day of the week, while elementary teachers work in a different school. (Time is provided for professional growth activities and planning in these schedules).

The Basic Mathematics Improvement Project, with a number of greatly different objectives, required a range of standardized and non-standardized instruments for its evaluation. Among the non-standardized instruments developed by the Evaluation Center were the following: (1) a *classroom teacher interview schedule* and a *mathematics improvement teacher interview schedule*, to gain insight into two views concerning the operation and effectiveness of the project; (2) a *materials rating scale* to determine objectives connected with uses of materials, the suitability of materials, the frequency of their use, and suggestions for additional materials for the project; (3) *attitude scales* for *teacher* and pupil assessment of individual students' attitudes toward mathematics; (4) a *mathematics department interview schedule*, to determine the effectiveness of the project as viewed by teachers of mathematics in the schools being served; (5) five forms of a *mathematics vocabulary and measurement achievement test*, to meaure attainment of the product objectives of the project; (6) a *supplemental data sheet* to obtain information concerning successes, problems, or recommendations of teachers in the areas of math instruction, professional growth, and program development; and (7) a *student information form* onto which all pertinent information concerning the student and his performance on relevant mathematics achievement tests could be entered, set up for key punching, and processed by computer. The standardized Science Research Associates Mathematics Achievement Test, forms A and B, were used to measure pupil achievement in the project.

The Reading Improvement Project, while similar in design to the Mathematics Improvement Project, is broader in scope and involves a greater number of students from the inner-city area. The project is designed to provide concentrated reading instruction, together with individualized assessment and guidance for disadvantaged students who are not reading at a level commensurate with their ability. The project provides services for over 2500 students from the fourth through the twelfth grades, with priority given to students from low-income families.

The project is staffed by reading specialists; one or more of these teachers are installed in each inner-city school in the project. The reading improvement teachers aid in diagnosing and recording each student's reading limitation and potential. Subsequently, these teachers help in grouping students according to the level and nature of their reading problems. Individual and group remedial instruction is then provided.

The reading improvement and classroom teachers work jointly to provide support for underachievers in reading. In secondary schools reading improvement laboratories have been established in conjunction with the project, so that students may work with the reading improvement teacher or by themselves to improve their skills.

Only non-standardized instruments developed by the Evaluation Center were used in the evaluation of the Reading Improvement Project. Among these instruments were the following: (1) a facilitant-restraint scale to determine factors of project operation which are affecting the reading improvement project; (2) a principal interview to gather information about the principal's perceptions of the effectiveness of the project; (3) a reading teacher interview guide, including both specific multiple-choice and open-ended questions to determine the reading teacher's perception of project efficacy; (4) a student interview focusing on reading interests; and (5) a teacher questionnaire for teachers in the intermediate and secondary grades.

The After-School Study Centers Project of the Columbus Public Schools attempts to creat an atmosphere conducive to study, in an area in which homes seldom provide intellectual and material resources to strengthen children's learning experiences. The project is aimed at upper elementary, junior high, and senior high students. It is housed in some twenty-four centers in schools throughout the inner-city and operates two (for elementary students) or three (for junior high and senior high students) nights a week, from suppertime until mid-evening. Use of the centers, regarded as a privilege rather than a right, hopefully will allow the student to improve his self-image, will stimulate him to value academic work more highly, and will motivate him to learn.

The centers are staffed by a supervising team composed of two competent teachers, cadet principals, or principals and a volunteer helping team. Volunteers are secured from a supporting community agency (for example, church, service organization, or university) and are chosen to provide capable persons who can serve as "models" for disadvantaged students.

Only non-standardized instrumentation was developed for the After School Study Center Project. A student interview was constructed to obtain information concerning the reasons for attendance or non-attendance at the centers, together with information on the usefulness to students of resources available there. In addition, a student perception scale was developed to gather data on the students' attitudes toward school, toward being assisted with their study habits, and toward the services and means of obtaining them which were present in the After School Study Centers Project. A teacher supervisor questionnaire was developed to determine what teachers did and thought about the operation of the centers. Finally, a similar questionnaire was developed for volunteer helpers to determine what activities volunteers engaged in during their work at the center and their perceptions of the successes and problems of the project.

The Health Centers Project is designed to expand and refine the medical and dental services presently operated by the school system and to bring selected services more closely in contact with disadvantaged students. On the assumption that good health is a prerequisite to a student's being able to benefit from learning experiences, the Columbus Public Schools have provided three health centers to serve the inner-city region. These centers provide physical examinations (including immunizations and inoculations), dental examinations (including follow-up, repair, and restoration), and referral of individual cases to medical-dental consultants, as well as providing the medical services component of the Pre-Kindergarten Project.

The Health Centers are staffed by one physician, two dentists (serving on a part-time basis), one nurse, and one dental hygienist. These persons and their equipment are housed in existing clinics in three junior high schools in the target area. Referrals to the centers are made by teachers, school nurses, and other medical personnel in the area.

Only non-standardized instruments were administered to the personnel of the Health Centers Project; the instruments developed included both accounting-type instruments and questionnaires and interview schedules. Included in the instruments which were created by the Evaluation Center staff were (1) a health center personnel interview, concerning staff perceptions of the needs of inner-city residents, the facilities of the project, and the emerging barriers to project success; (2) a data collection sheet for the project, for recording information concerning the movement of persons through the centers, and the type and disposition of their treatment; (3) a parent interview, attempting to gather information concerning parents' perceptions of treatment by the health center staff and the degree and kind of communication between staff and patients; (4) a school nurse questionnaire, focusing on the involvement with and assessment of the project; and (5) a teacher awareness questionnaire and (6) principal interview, both attempting to gather information on the use of the centers by teachers and principals and the impact the operation of the centers had on the health problems of students and the operational problems of schools and their staff members.

III. A CRITICAL LOOK AT EVALUATION AND ITS INTENT

Evaluation to serve decision-making requirements must be valid, reliable, timely, credible, and accessible. To what extent has the Columbus evaluation effort achieved these objectives? Unfortunately no rigorous and thorough study of evaluation has been made during the course of the Columbus project. However, a number of judgments can be made about the impact of the process and perhaps about its effectiveness in reaching these objectives. Validity and reliability are criteria which all information gathering efforts attempt to meet. The objectives which separate the concept of evaluation from inquiry in a more general sense, and certainly from "basic" research, are the objectives of timeliness, credibility, and accessibility. It is the latter three objectives which are especially important to those attempting evaluation in an ongoing school program, for attainment of these objectives is necessary in order for evaluation to be useful to decision makers operating with time constraints.

The validity of evaluation efforts is difficult to assess, for only at times does an explicit theoretical framework exist on the basis of which one can make an evaluation design and appropriate instrumentation. More often, validity is determined (as it was in Columbus) by reference to those variables which decision makers find of interest. Because of the selective concerns of various decision makers, only some of the many variables which could have been investigated were focused upon by evaluators in theColumbus project. When variables were focused upon, literature searches were undertaken to provide some foundation for instruments developed within the Center. Standardized tests were assessed for their relevance to the objectives of the projects and for their usefulness in providing appropriate information quickly. At present, little evidence exists to prove whether these tests were relevant, beyond the belief of decision makers. Whether they were in fact appropriate to the Columbus Title I effort (valid for its disadvantaged students) may not be known until some time in the future. The validity of process instruments (the type most often constructed at the Evaluation Center) is even more difficult to ascertain, for the variables--their content and form--which these instruments attempted to measure were constantly evolving throughout the course of the project. Process variables investigated were those which seemed at the moment to be of importance; their importance in other projects or in this project at a later date cannot be known. While it might be useful to attempt to relate process variables to, for example, organization theory, such determinations must be made after the fact for Columbus projects.

Validity, in terms of the variables which interested decision makers, appears to be high. Because of the evolution of the project, however, the objectives and processes for attaining them have changed, and the only thing which can be said is that the evaluation effort appeared to be valid at a particular moment.

The reliability of the evaluation has not been studied formally. Overall, there is no way to determine the likelihood that similar results would be obtained from a similar evaluation conducted a second time--except as subjective probability tells us that it is likely. Often, when similar instruments were used more than once, results which were consistent with earlier results appeared. However, using these findings as an indicator of reliability must be done cautiously, for if instruments were used more than once, it was likely that they were in a different, changed form the second time. Few instruments remained unaltered; rather, they evolved as did the project. Many instruments were employed only once, on the other hand, to gather data which appeared relevant at one moment, but appeared less useful at later phases of the project. In addition, when instruments were used over a long period of time, or over wide groups, a setting variable appeared to be acting which might have influenced the reliability of findings. Because samples of respondents used in the evaluation were such tightly-knit groups with all members being in close contact, respondents to similar instruments tended to give collusive information. Group members tended to communicate with one another about the types of questions asked by evaluators, and often preconceived notions about questions (and preconceived answers) were awaiting the evaluator when he came to work with a respondent. Such collusion presents a great problem if incorrect assessments prevailed in the minds of respondents. The fact that evaluators were unable to determine reliability of findings was constantly pointed out to decision makers involved in the evaluation effort; the degree to which this precautionary statement was believed, or played a role in later decisions, is as yet undetermined.

Achievement of the objective of credibility can be accomplished by two means. First, the decision maker himself may be familiar with the area under investigation and can determine (on the basis of his technical expertise) the credibility of findings presented by another person (i. e., the evaluator). On the other hand, if the decision maker is not intimately familiar with the area under investigation, he often forms notions of credibility based on his assessment of the individuals presenting information to him. Both of these means of building credibility have been used in Columbus. In at least one project decision makers were doubtful of the credibility of evaluators and only begrudgingly granted the trustworthiness of information. This situation came about through an interaction of familiarity with information on the part of decision makers and an uncertainty on the part of both decision makers and evaluators as to the exact confines of each role. In this specific project, a gradual deterioration of relations between the evaluation staff and project decision makers occurred. Presently, new methods are being used in an attempt to rebuild credibility in the eyes of these decision makers. (However, in such a case it is often necessary that both sides give a bit of ground.) To rebuild credibility, the design and instrumentation for the evaluation effort in the coming year is being geared particularly closely to the needs of decision makers. At the same time, the precise role of evaluators is being defined, so that little contention about who ought to do what can occur.

By and large, sound working relationships have been achieved between evaluators and decision makers of individual projects. It would appear that, for credibility to be achieved, the evaluator must act in a professional manner, make his commentary on findings and analysis precise, and attempt to make his values explicit. It is important for the decision maker, conversely, to be precise about his objectives and certain about the type of information he believes useful. In this way, the evaluation effort may be less threatening and more reliable and valid, and evaluators and decision makers may work together closely.

The requirements for credible and accessible information by all decision makers in the project tend to interact to produce possible fluctuations in reliability. In the Columbus project, for example, the orientation of the evaluation staff was predominently to top-level decision makers in projects. However, evaluators themselves often had to obtain information from lower-level decision makers or non-participants in the formal project decision-making process. Because information obtained from respondents often was not accessible to them after it had been analyzed and reported, lower-level decision makers and non-participants tended to view the evaluation as not being particularly credible, nor its findings accessible for their needs. Hence, these persons became disenchanted with the progress of the evaluation, and informal feedback sometimes was necessary to insure their continuing cooperation.

Timeliness, one of the most important objectives that evaluation must meet, has been only partially met in the evaluation thus far. Certainly it must be a common complaint among project administrators of all sorts that social scientists tend to be of limited use in obtaining information to aid in analyzing pressing problems. From an administrator's viewpoint, it often appears that by the time social scientists have defined the problem, debated methodology, proceeded with the collection of data in an acceptable and "clean" manner, analyzed the data, and reported their findings the problem which motivated this analysis has long since been solved, in a judgmental matter, by a decision maker pressed for time.

The concept of evaluation employed in Columbus is an attempt to circumvent this objection. Concerns about absolute validity and exceedingly high reliability have been slightly forsaken so as to increase the timeliness of and accessibility of information. It is important to recognize that any evaluation effort must come to terms with all the objectives of sound evaluation and achieve a mix reflecting the needs of the situation. Seldom can evaluation efforts achieve all of the objectives to the fullest degree.

Timeliness is a major concern in the Columbus evaluation effort. The best that can be said thus far, however, is that data provided to decision makers were provided periodically; the data were not provided continuously, nor always at the exact moment when they were most needed. Different modes of returning information to decision makers were employed, resulting in different degrees of timeliness. Periodic feedback sessions were held rather often for each project; however, in most instances an average of two months transpired between the time instruments were made ready and administered and the time that a complete report was given to decision makers. In the time between, informal means of transmitting information were employed.

Quite often decision makers would remark when confronted with a piece of information "I knew that." However, the interpretation of this is open to debate: it may have been that decision makers "knew" in an intuitive fashion; it may have been that they knew on the basis of earlier, informal contact with evaluators; or it may have been that they simply were defensive or anxious to get on with other things. The formal feedback sessions were only somewhat timely; few crucial decisions were made on the basis of information first presented there.

Some sort of determination must be made concerning the rough trade-off of timeliness for achievement of other objectives of sound evaluation prior to creating the evaluation design for a project. Timeliness can be achieved, of course, but at a cost in the attainment of other objectives (though it is impossible to say exactly how much of one will obviate some of another). The determination of this trade-off should not be made by decision makers alone, for often the evaluation staff will believe they have a responsibility for the validity and reliability of information and will be willing to forego only a limited time before becoming adamant that these other objectives be considered.

Accessibility as an objective of sound evaluation presents an exceptionally sensitive issue. A policy about accessibility must be formed by decision makers concerning those persons who "have a need to know."

In Columbus certain respondents to evaluators believed that they ought to have some feedback (some accessibility to the information) on the results of the evaluation. The determination of how to handle this claim cannot be made by evaluators; it must be made by decision makers responsible for the project.

An even more sensitive problem is that of what might happen should unknown persons obtain evaluation information. Public school personnel in the Columbus projects were acutely sensitive to the desires of interest groups in the city to obtain information about the ongoing Title I program. There was often the concern in writing reports (which the Center's philosophy implied ought to go to decision makers) that nothing be said which, if it "leaked out," could damage the project. Obviously, determining who has a need to know must be done both for persons within and outside the institutional context in which evaluation takes place.

Some mechanics of communication are involved in achieving accessibility of evaluation information. Reports need to be written in such a way that persons with different backgrounds can read, comprehend, and utilize the information contained in them. (At the same time, of course, it is important for evaluators to insert technical considerations and disclaimers about their data and methods.) Creation of an accessible report and the use of informal communication channels were the primary means for achieving accessibility of evaluation information in Columbus. Other means of providing access, such as the use of data banks and inventories, were not attempted due to the shortage of time, money, knowledge about who had a need to know what types of information, and knowledge of how useful such mechanisms would be. The present effort appears to be reasonably accessible to those persons who, prior to the beginning of the evaluation effort, were identified as having a need to know--the top-level decision makers.

A concern of evaluation not explicit in the previously-discussed criteria is sometimes considered at the Evaluation Center. This concern is comprehensiveness--the scope and depth of data which will be collected and analyses which will be performed before information is presented to decision makers. Decision makers are often quite explicit about what they need. They are concerned about the nature of a particular problem, the nature of a certain variable, or the interaction of specific elements in the system that they are operating. Often, however, it is useful and perhaps significant (in a practical political sense) to relate specific problems to long-run or "spillover" effects which problems and potential solutions create and to the costs of eliminating these effects. Values espoused in the process of rational decision making can be narrow or broad on the issue of comprehensiveness. It is important to determine the scope of the evaluation effort and the point of view which evaluators shall take in the gathering, analysis, and reporting of date, for the simple reason that efficiency is achieved more easily if these values are made explicit.

As mentioned earlier, the Columbus evaluation is focused on the parts of a system; it does not represent a systems analysis of the projects, and only a partial analysis of each individual project. The evaluation is concerned mainly with information relevant to problems. If a total system view appears valuable, the decision maker must choose to take it and realize that taking such a view is costly (especially since the methodology for gathering information from a comprehensive viewpoint is vague). In order to take a comprehensive point of view, the scope of comprehensiveness must be defined by decision makers, so that the evaluation effort can be precisely focused and problems of methodology identified. As this is done, the relation of comprehensiveness to other objectives for a strong evaluation can be more firmly defined.

In summary, the Columbus evaluation effort has been valid to the degree that is necessary for decision makers in the project. It has been reliable to the degree possible given constraints of time and the requirement that instruments constantly change. It has been reasonably timely; by and large it has been credible; and it has proved accessible to those who were defined as having a need to know. A formal evaluation of evaluation must be undertaken to provide a description of its exact impact on the behavior of decision makers in a project. For the moment, it is hoped this informal view and judgmental criticism of evaluation will allow others to begin to make new thrusts in this changing field.

Part II

THE PITTSBURGH EVALUATION PROGRAM

THE PITTSBURGH EVALUATION MODEL

Background and Context

In Pittsburgh Public Schools a model evaluation plan is being developed under a contract with the United States Office of Education and in conjunction with the evaluation of Title I programs. The proposal for the development of model criteria and procedures in Pittsburgh set forth the following premises:

1. That evaluation in the public school setting is a process for program improvement as well as for program assessment
2. That maximizing the involvement of program personnel (field staff and administrative staff) in the process of evaluation fosters staff commitment to program improvement and promotes desired change in staff behavior
3. That the nondirective role for evaluation personnel is effective in promoting program improvement

In addition, the following conditions or factors in the Pittsburgh setting have contributed to shaping the Pittsburgh Model:

1. Lack of adequate pre-implementation planning and evaluation for Title I programs
2. Anticipation of eventual support for evaluation of other educational programs currently operating within the system

3. Proliferation of Title I programs (28 during the first year of model development), which has made necessary some compromises between the optimum and the practical in developing evaluation procedures

Out of the foregoing considerations and the experience of model building, the following features of the Pittsburgh Model have emerged:

1. The purpose of program evaluation is seen as that of providing information requisite to program development and stabilization and for valid program assessment to those responsible for decisions to change the program.
2. Evaluation and decision-making functions are seen as separate.
3. Educational programs are viewed as discrete subsystems of the total school system.
4. The audience for evaluation information is seen as primarily the decision makers within each program. The exceptions to this rule are at points of inter-program contact (e.g. are separate programs compatible?) and comparison (e.g. which is most efficient?).[1] Decision-making responsibility is

---

[1] In order to satisfy the requirements of the Elementary and Secondary Education Act of 1965 and of the state and federal agencies responsible under the law, product assessment not always consistent with the Pittsburgh model is carried on concurrently with model activities. Product evaluations not consistent with the model are reported annually, whereas evaluation reports under the model are not tied to a fixed time schedule.

presumed to reside in all strata of program staff; i.e., all members of program staff, from teachers on up through the supervisory and administrative ranks, are potential program developers.

5. The focus of evaluation under the model--the evaluation criteria and their sequence--reflects an interpretation of the information input necessary for program improvement and assessment in the public school setting. Program development is seen as a spiraling process, with much recycling of change activity as a program evolves.
6. The procedures of evaluation under the model--the specifications for how evaluation purposes are to be carried out--are engineered to encourage the involvement and commitment of teachers and other categories of field staff to insure acceptance and use of evaluation findings.

The Focus of Evaluation

The first concern of the model building effort in Pittsburgh is that of determining what is to be evaluated. Guidelines have been adopted which relate evaluation to the spiral of program development, specifying both the sequence of evaluation interest over time and the scope of interest in point of time. An explanation of the guidelines and their underlying rationale is presented in the following section.

The Guidelines: Sequence

Guidelines relevant to sequence in evaluation are represented in Figure I. Under these guidelines, evaluation purposes or objectives are seen as related to the decision problems of program developers; these decision problems, in turn, are seen as being determined by the status of the program in the process of program development. For convenience in formulating a generalized plan of sequence, the development process is viewed as being segmented into four levels or stages, and a set of evaluation objectives and criteria is associated with each stage. In Figure I, the four evaluation stages are listed in Column I, and a set of evaluation objectives for each stage in Column II.

Under the guidelines, each evaluation objective is implemented by the evaluation staff through the observation of specified phenomena and the application of specified criteria in making evaluation judgments. To the extent that it is possible at this time to generalize, the guidelines specify the sources of standards for the criteria. In Figure I, the observations and criteria associated with each objective are listed in Columns III and IV respectively. The standards sources for specified criteria are listed in Column V. The development problems for which the evaluation produces information are listed for each stage in Column VI.

Stages of Program Evaluation I	Evaluation Objectives II	Phenomena Observed III	Criteria for Evaluation Judgments IV	Standards Sources V	Development Problems for Decision VI
STAGE I	To assess viability of program design	Public consensus of program staff in response to questions about program dimensions (program design)	Comprehensiveness of program design Face validity of program design	Pittsburgh taxonomy of program dimensions Consultant employed by evaluation staff	Should resources be reallocated within the system? Should the operating program be terminated? Should resources be reallocated within the program (1) to extend the program design or (2) to alter the program design?
	To assess feasibility of program design	Perceptions and opinions of field staff about reciprocal effects of program and program environment	Compatibility of program with program environment	Hierarchy of system objectives implicit in opinions of field staff	
STAGE II	To assess implementation of program and To assess validity of program assumptions	Public consensus of program staff in response to questions about program dimensions (program design) Status of antecedent and process dimensions in operating program	Congruence between operating program and program design in antecedent and process dimensions		Should resources be reallocated to alter the program design and/or to effect operational adjustment (broaden and internalize the public consensus)?
STAGE III	To maintain fidelity of operating program with program design	Public consensus of program staff in response to questions about program dimensions (program design) Status of antecedent and process dimensions in operating program	Congruence between operating program and program design in antecedent and process dimensions		Should resources be reallocated to effect operational adjustment (broaden and internalize the public consensus)?
	To assess effectiveness of operating program and validity of program assumptions	Status of outcome dimensions in operating program	Congruence between operating program and program design in outcome dimensions	Program design relative to criteria for objectives	Should resources be reallocated to alter or restructure the program design?
STAGE IV	To maintain fidelity of operating program with program design	Public consensus of program staff in response to questions about program dimensions (program design) Status of operating program in all dimensions	Congruence between operating program and program design in all dimensions		Should resources be reallocated to effect operational adjustment (broaden and internalize the public consensus)?
	To assess efficiency of operating program and To assess economy of operating program	Time cost of operating program Dollar cost of operating program	Cost relative to effectiveness Cost relative to efficiency	System values System values	Should the program be terminated? Should the program design be altered? restructured?

Figure I
Evaluation Guidelines: Sequence

Stage I. The objectives for evaluation in Stage I are to assess the viability and the feasibility of the program design under the condition that the program is already operational; the paramount question--which definition, or design, of the program to assess--is resolved by the guidelines, which provide for generating a dynamic baseline definition, or working definition, of the program by means of public interviews with representative groups of program staff.[2] This observed public consensus is then judged for viability by applying the two criteria of comprehensiveness and face validity.

The standard used for the criterion of comprehensiveness has been developed in Pittsburgh through systematic analysis of the definitions or designs of all Title I programs. This standard is embodied in a comprehensive list of program elements and is presented in Figure II as the Taxonomy of Program Dimensions. As shown in

---

[2] After a Title I program has been implemented in Pittsburgh, there are usually at least three designs of the program in existence: one is represented by the project proposal, at least one other exists in the thinking of program leaders, and a third is reflected by what actually happens in the operation of the program. Experience suggests that, at any given time, the actual number of designs in existence for a given program is proportional, inversely, to the quality of the intrastaff communication achieved for that program. The rationale for the consensus definition is as follows: it provides an expedient focus for program development activity aimed at modifying and internalizing program goals because (1) it reflects many of the divergent views held by the program staff, while at the same time, (2) it is largely shaped by the thinking of program leaders whose influence in the process of deriving the definition may act as a stimulus for internalization.

GENERAL DESCRIPTION	OUTCOMES	ANTECEDENTS	PROCESS
I. Overall Statement of Objectives and Rationale of the Program II. Description of Scope A. Number of Pupils and/or Number of Classes and Schools Involved B. Grades or Ages of Participants C. General Description of Staff	I. Major Objectives A. Terminal Objectives[1] B. Ultimate Objectives[2] II. Enabling Objectives[3] III. Other Benefits[4] IV. Criteria for Successful Completion of or Removal from the Program	I. Students A. Selection Criteria[5] B. Entering Behaviors[6] II. Staff A. Program Personnel by Specific Positions B. Qualifications for Specific Positions 1. Professional 2. Personal III. Support A. Administrative Support B. Human Resources C. Media D. Facilities IV. Time Constraints	I. Participant Activities A. Enumeration of Activities B. Estimate of Time Spent on Each C. Media Used to Carry out Activities D. Activities Related to Objectives II. Staff Activities A. Functions and Duties for Specific Positions B. Intra-staff Communication and Coordination C. Communication between Program Staff and Others and its purpose

[1] Behaviors exhibited by participants at the end of the program which demonstrate successful completion of the program

[2] Long range goals of the program, objectives to which the program hopefully contributes, but for which it does not have sole responsibility

[3] The skills, attitudes, and information which students must acquire during the program to insure the accomplishment of the major objectives

[4] Benefits expected to accrue to other than program participants as a result of the program

[5] Criteria that are used to determine who shall participate in the program

[6] Characteristics of participants (other than selection characteristics) which are related to performance in the program

Figure II
Taxonomy of Program Dimensions

Figure II, program elements are classified into four broad categories. Three of these--outcomes, antecedents, and process--are regarded as necessary and sufficient for program design. The educational program is conceived as a dynamic input-output system; in keeping with this concept, channels and procedures for communication are seen as essential elements of program design.

The basic questions relative to comprehensiveness of the program definition are (1) "Is there specific program information for each dimension in the taxonomy?" and (2) "Are program dimensions spelled out in acceptable form, i.e., are objectives stated in terms which describe behavior, conditions of behavior, and standards or criteria?" In judging a design for comprehensiveness in the first or second cycles of evaluation,[3] it is common in the Pittsburgh experience to find that many of the dimensions are either not specified, not complete, or not in desired form (e.g. objectives are not stated in behavioral terms). The implication for decision makers in these instances is that curriculum development work or other activity to extend and/or modify the program definition should be initiated.

The second criterion for viability of the program definition is that of face validity, or the logic (reasonableness) of the functional relationships implicit in the definition. The standard source suggested by the model for application of this criterion is a qualified consultant employed by the evaluation staff.[4] In judgments relative to face validity reside implications for modification of program design.

---

[3] A description of the evaluation cycle appears in a later section of this paper.

To implement the second objective of Stage I evaluation, that is, to assess the feasibility of the program definition by applying the criterion of compatibility, a second set of observations is collected. These observations relate to the reciprocal effects of the program, as currently designed, and the program environment. The question to be answered is "Does this program conflict with other programs or with the school system as a whole in regard to the use of student time, staff time, facilities, or media?"[5]

In the Pittsburgh system, information on which to judge compatibility is not readily available at present. The current practice, therefore, is to assess the perceptions and opinions of field personnel, accepting as a standard the hierarchy of system objectives which is implicit in those opinions. Judgments arrived at in this manner are indeed gross; when unquestionably negative, they lead to questions for

---

[4] A checklist useful for assessing face validity is presented in Figure III. These statements are derived from "Criteria for Stage I Evaluation Judgments" presented in Figure 6 which guides the Pittsburgh evaluation staff.

[5] The basic questions which can help determine compatibility are presented in Figure IV.

	Yes	No
1. The terminal objectives tell what the student does to signify success in the program.	___	___
2. The terminal objectives are clearly related to day-to-day activities of the program.	___	___
3. The enabling objectives are clearly related to the terminal objectives.	___	___
4. The entering behaviors are consistent with the selection criteria.	___	___
5. The entering behaviors are linked to the program's objectives.	___	___
6. The media are clearly related to activities and are sufficient for the essential activities.	___	___
7. The activities are related to the objectives. There is at least one activity for each objective.	___	___
8. Time resources are realistically related to the objectives. There is sufficient time for each specified activity.	___	___
9. Staff qualifications and characteristics are adequate for the functions defined. If not, in-service activities are spelled out to make them adequate.	___	___
10. All functions necessary to serve the objectives are included.	___	___
11. The functions for each group are clearly related to program objectives.	___	___
12. Staff duties are clearly related to staff functions.	___	___
13. Staff duties are clearly defined.	___	___
14. There are duties for each function.	___	___
15. Intra-staff activities are adequate for support of program operation or program objectives.	___	___
16. Communications channels outside the program are related to support needs.	___	___

Figure III
Checklist for Face Validity

1. Is sufficient time available for student participation in the program?
   What activity does the student give up in order to participate in the program?
   Does this reallocation of student time result in sacrifice to other objectives of the school program?
   Does it have an effect on the operation and/or goal attainment of this program?

2. Is sufficient time available for participation by the program staff and cooperating personnel?
   What activities do staff or cooperating non-program personnel sacrifice in order to participate in the program?
   Does this reallocation of their time result in a sacrifice to other objectives of the school program? How does it affect this program?

3. Are facilities and media now available to the program? If not, have plans been made to provide them?
   Is allocation of facilities and/or media to this program resulting in sacrifice of other objectives of the school program?
   Is this program affected by the manner in which facilities and media are allocated?

4. Are the gains for students anticipated by this program equal to, less than, or greater than possible sacrifices in other educational objectives of the school program?

Figure IV
Basic Interview Questions, Program Compatibility

decision making relating to reallocation of resources within the system, program termination, and program modification.

Stage II. The objectives for evaluation in Stage II are to assess both the implementation of the program and the validity of the program assumptions. The initial focus is the current status of the working definition, as reflected by the current public consensus of program personnel. The second set of observations for Stage II evaluation relate to the operational reality of the program in the antecedent and process dimensions.

A single criterion is applied--the congruence or incongruence of current program design with current program operation. This congruence or lack of congruence reflects both on the quality of implementation and/or the validity of program assumptions. If program operation does not accord with intent (as reflected in the public consensus) two, and only two, inferences are possible: (1) functional relationships implicit in the program definition--the program assumptions--are invalid and/or (2) the program has not been communicated and/or internalized among the specified personnel. Implications from these findings relate to action for program redesign and/or for operational adjustment through broadening and internalizing the current consensus.

The guidelines do not yet specify the standard for congruence. The question to be explored is, "How much discrepancy between program design and program operation should be regarded as inevitable?"

Perhaps the standard which evolves will be a ranking of antecedent and process dimensions in terms of the relative importance of discrepancy in each dimension.

Stage III. The observations and judgments which implement Stage II objectives serve another purpose at Stages III and IV. At these higher stages in the sequence, evaluation serves both to monitor the quality of implementation, and as stated in the guidelines, to maintain the fidelity of the operating program with the program design. Thus, at Stage III the working definition is compared with the observed status of antecedent and process dimensions in the operating program. Any resulting implications for change activity relate to operational adjustment through improved communication of the program.

The other purpose of evaluation at Stage III is assessment. Both the effectiveness of the operating program and the validity of program assumptions are assessed in one set of operations--the observation of program outcomes and comparison of these for congruence with outcomes specified in the design. The source of standards for judging congruence is identified by the guidelines as the program's own design (working definition). As seen in the Taxonomy (Figure II) one necessary element of program design is the specification of criteria for outcomes.

Inferences inherent in negative assessment at Stage III relate to validity of the program design. Instability as an explanation of failure has been ruled out at this stage by both the accomplishment of evaluation objectives at Stage II and the subsequent quality control function of evaluation. Therefore, when it is found that the operating program is not achieving its objectives, the validity of functional relationships implicit in the program definition is called into question. The implication for program leaders is change activity directed toward altering or restructuring the program design.

Stage IV. The quality control function of evaluation continues in this stage and is represented in the guidelines as the first objective for Stage IV evaluation--to maintain the fidelity of the operating program with the program design. As in Stage III, this objective is implemented by first observing both the current working definition of the program and the status of the operating program and then applying the criterion of congruence. Implications from negative evaluation relate to operational adjustment through improved communication of the program.

The unique objectives of evaluation in Stage IV are to assess the efficiency and the economy of the operating program. The observation relevant to assessing the efficiency of the program is its cost in terms of student or participant time (one aspect of the antecedent dimension). The criterion of cost relative to effectiveness is applied, invoking standards derived from and for the entire school system.[6]

---

[6] These standards do not yet exist in the Pittsburgh system.

Implications from negative assessment in Stage IV relate to decisions for program termination or program modification.

The Guidelines: The Evaluation Cycle

The guidelines for sequence are purposefully inexplicit about the time dimension in evaluation. The range of evaluation interest in relation to periods of time during program evaluation is clarified in the explanation of the evaluation cycle--the mechanism for adapting the evaluation sequence to the rate of change in the program. Under the guidelines, a cycle of evaluation consists of those evaluation activities essential to implementing a set of evaluation objectives and, in addition, a monitoring of program change at the end of the cycle. The evaluation objectives which define an evaluation cycle are not identical, necessarily, with any single group of objectives that defines an evaluation stage (see Figure I).

After the first cycle, each succeeding evaluation cycle is defined in part by the necessity of evaluating in the light of changes in the program. Thus, as program development activity recycles, so also does evaluation activity. As successive evaluation cycles adopt new objectives moving ahead through the sequence of evaluation stages, they also include some of the objectives and associated activities of earlier stages. After Cycle I, an evaluation cycle is likely to be broader in its scope of interest than a single evaluation stage.

Criteria for Definition of the Evaluation Cycle. Experience with the interaction of evaluation activity and program change currently is insufficient to suggest more than general guidelines for cycle definition. These are as follows:

1. Negative Evaluation--When a negative evaluation judgment is reached during one cycle of evaluation, the associated criteria and activities are included again in the next cycle of evaluation. New objectives associated with a higher stage of evaluation are incorporated as long as consistent with the actual progress of program change. There is always a time delay between evaluation judgments directed toward program change and actual program change. Furthermore, pursuant to negative judgment at a given level of program evaluation, program development must recycle at this level and possibly also at earlier levels before "catching up." The decision to define a new evaluation cycle in terms of the next higher level of evaluation is based on expectation that program recycling will be rapid enough for sufficient progress within the time constraints imposed by the structure of a new evaluation cycle. Considerations which affect the expectation of adequate program progress are as follows:

a. Observation of the change decisions and plans of program leaders and judgments of the efficacy of these plans by those members of the evaluation staff most knowledgeable about the program. (Is the impetus and know-how for change activity adequate? Is the available time adequate?)

b. Nature of the evaluation observations needed at the higher level of evaluation and time constraints affecting the collection of such observations. (Can program design changes be completed and monitored in time to determine the source of observations and to make collections within the duration of the ensuing cycle?)

2. Positive Evaluation--When positive, or predominantly positive, judgments are reached during a cycle of evaluation, the ensuing cycle is defined primarily by objectives and activities associated with the next higher stage of evaluation.

The Procedures of Evaluation

Guidelines for sequence and cycling define the focus of evaluation. Equally important in the Pittsburgh Evaluation Model are the procedures for implementing these guidelines. The procedures currently employed for making and reporting the specified observations and judgments are described below. For the most part, these apply for Stage I evaluation (project development activity is presently focused on Stage II procedures), although their potential for use in subsequent stages is apparent. The premises of maximum participation of program personnel and nondirectiveness on the part of evaluation personnel have played a prominent role in defining the evaluation procedures herein described.

The Program Definition Meeting: Rationale

At the start of evaluation, Title I programs are ill-defined and poorly communicated. Lack of agreement on program objectives is reflected in the distinctiveness of the goal descriptions from different sources. Typically, at least three distinct sets of goals are presented: one set by the project proposal; a second set by program leaders; the third and additional sets of goals, by implication, in the field operation of the program. Deficiencies of program design are evidenced in lack of specificity regarding the antecedent conditions and procedures for reaching objectives and in the vagueness of the objectives themselves.

Whereas the confusion of goals and deficiencies of design suggest the emphasis for evaluation in the initial stages, they also represent a substantial procedural barrier: whose objectives and whose definition of the program do we evaluate?

The solution is provided by the Program Definition Meeting, a procedure of structured group interview for eliciting and recording the public consensus of the entire program staff (or a representative sample) on questions relative to program dimensions. This procedure encourages maximum participation of program personnel in the process of deriving a baseline, or working definition, of the program. The product of this meeting--the working definition--reflects many of the divergent views held by program staff, while at the same time it is largely shaped by the thinking of program leaders whose influence may act as a stimulus for internalization. Thus, it is a convenient focus for program development activity aimed at modifying and internalizing program goals.

The Program Definition Meeting: Description

A. Purpose

1. To generate a definition of the program which
    a. Is as complete as possible in terms of the Pittsburgh Taxonomy of Program Dimensions (See Figure I)
    b. Reflects divergent opinion within the program staff
2. To encourage communication between levels and within levels of program staff and to increase knowledge about the program
3. To encourage acceptance of the value of evaluation
4. To maintain a nondirective role for evaluation staff

B. Participants

1. Program staff
    a. When possible, the entire program staff, including the chief program administrator and his superior in the administrative ranks, is invited to the meetings.

   b. When participation of total staff is not possible, a stratified sample of staff is invited. Variables taken into account are those which relate to variability of program goals such as (1) length of service in the program, in the school, and in the profession; (2) the size of the school; (3) socio-economic characteristics of the school; and (4) function in the program.

2. Program staff participants enter the meeting with varying degrees of interest in and knowledge about the program. Many, especially at higher levels, are apprehensive about the effects of evaluation. Many at lower levels are reluctant to express divergent opinion in the presence of program leaders.

3. Evaluation Staff

   a. Discussion leaders: Members of the field research staff with leadership experience and/or training in theory of group process and in techniques of discussion leadership play a major role in the definition meeting. These leaders are briefed one or two days before the meeting by the program evaluator.

   b. Status leaders: The Director of Research and/or the Coordinator of Evaluation are present at the meeting.

c. Program evaluator: The member of the evaluation staff with major responsibility for evaluation of the program is involved before, during, and after the meeting.

C. Other Resources

1. A large meeting room with space and facilities for subdivision of the participants into a series of small (maximum of 10 participants) discussion groups is used. The maximum number of small groups accommodated is six. Thus, the total number of participants per meeting is limited to 60.

2. One half-day is the limit of time available to members of the instructional staff for participation in the Program Definition Meeting. Definition meetings adapted to this constraint are conducted for a three-hour morning or afternoon session.

D. The Process

1. Planning and coordinating the meeting is the responsibility of the program evaluator. Prior to the meeting he completes the following plans and preparations:

   a. Arrangement of details of date, time, and location of meeting and selection and inviting of participants. In all of these activities the program evaluator works closely with the program administrator or other program leader.

b. Preparation and distribution of the discussion agenda. Dimensions of the program as specified in the Taxonomy of Program Dimensions are interpreted to program personnel in the form of questions about their program. The program evaluator phrases these questions in terms familiar to program personnel. Copies of the discussion agenda, or interview schedule, are distributed to the participants in advance of the meeting.

c. Assignment of participants and discussion leaders to groups. Specific assignment of participants serves several purposes:

(1) To represent within each group those variables--function, attitudes, status, length of experience, physical location--which relate to variability in program goals

(2) To set the stage for maximum input by each individual in the group through strategic placement of authority figures and prevention of interpersonal status barriers

(3) To limit the size of groups to a maximum of ten members each

d. Briefing of discussion leaders. At least one day before the meeting, the program evaluator conducts a briefing for discussion leaders during which the following kinds of information are presented:

(1) History of the program's development and implementation

(2) Rationale and general description of the program

(3) History of evaluation of the program

(4) Interpretation of the interview schedule

(5) Composition of individual discussion groups, including identification of variables relevant to the discussion process known by the evaluator to be present within the groups--role or value conflicts, status sensitivities, attitudes of program personnel, etc.

(6) Details of physical arrangements, time allocation, and other administrative matters

2. The meeting is opened by status leaders. Leaders from outside the evaluation staff are encouraged to make opening remarks and to lend their support to the goals of the meeting. The status leader representing the evaluation staff explains the nature and purpose of the meeting, including its place in the evaluation process, and stresses the value and importance of individual contributions in the process of program definition. No more than fifteen minutes is devoted to this part of the program.

3. The participants are assigned to small groups for the interviewing. The discussion leader for each group explains again the purpose of the group session and the details of produre:
   a. He stresses the equality of participants in the process and the value of every contribution.
   b. He describes the method of circular response, the procedure for reaching the goal of equai and/or maximum individual participation.
   c. The leader may use a technique such as introduction of individual group members as a means of establishing a favorable climate for the work of the group.
4. The discussion leader uses the interview schedule and the method of circular response to interview the group. The leader continually interprets, feeds back, and summarizes the responses of participants in order to record group consensus for each question. When consensus is not possible, division of opinion is recorded. (The leader may appoint a recorder to supplement his own note-taking.) The interview continues for two hours or more.
5. The program evaluator monitors the discussion groups and decides when the group interview session can be terminated.

At his signal, participants reconvene as a unified group for concluding statements by the program evaluator or the status leader. The leader describes the use to be made of the group interview data and promises feedback to the participants in the form of a formal definition of the program.

6. The program evaluator conducts a post-meeting session of group leaders for discussion and analysis of the interview processes and the resulting products. The evaluator gains insights which help him in the analysis and synthesis of the group interview data and the preparation of a formal definition of the program.
7. The evaluator prepares a formal definition of the program which is then distributed to all participants or to all members of the program staff.

Stage I Panel Meeting: Rationale

A danger inherent in judgmental evaluation is subjectivity in the choice and application of judgment criteria. Alternatives for either avoiding or mitigating the effects of this danger are (1) to confine evaluation to information gathering without judgment, or (2) to employ only criteria identified by the interested parties. The Pittsburgh Evaluation Model rejects these alternatives as inefficient for the purpose of stimulating program improvement.

To preserve the judgmental function of evaluation while minimizing the inherent danger, the Pittsburgh Model relies on two strategies. First, evaluation criteria are predetermined, universally applicable (that is, applicable to all programs evaluated), and external to the individuals responsible for applying them. Secondly, on the presumption of greater safety in numbers, the model employs the panel meeting as the procedure for making evaluation judgments.

Stage I Panel Meeting: Description

A. Purpose

1. To make judgments of the program definition by applying the Stage I criteria
2. To encourage acceptance and use of findings by program decision makers
3. To maintain a nondirective role for evaluation staff

B. Participants--Members of the Panel

1. The program administrator, who may enter the panel meeting somewhat apprehensive about the role and effects of evaluation
2. The program evaluator, the member of the evaluation staff with major responsibility for evaluation of the program who is involved before, during, and after the meeting
3. A consultant employed by the evaluation staff who is a content

specialist in the area of the program's major emphasis

4. A resource person from the evaluation staff who is technically competent in matters of instructional objectives and program design
5. A status leader from the evaluation staff, usually the Coordinator of Evaluation

C. Other Resources

1. Facilities and Equipment

The desired atmosphere is best achieved in a small, comfortable room containing a round table with capacity for seating five persons. A tape recorder operates throughout the meeting to record the deliberations and findings of the panel for subsequent analysis.

2. Materials

Before the meeting, copies of the Program Definition, product of the Program Definition Meeting, and the Taxonomy of Program Dimensions (See Figure II) are sent to each participant. The program administrator and the consultant are given, in addition, materials which describe the purposes and procedures of the panel meeting (See Figure V, Guidelines for Stage I Judgments). During the meeting, the panel moderator is guided by these procedures and by the Criteria for Stage I Evaluation Judgments (See Figure VI).

1. Judgments of the Program Definition will be made by a panel consisting of the program evaluator, the program director, one consultant employed by the Office of Research, and one resource person from the Office of Research, with the Coordinator of Evaluation acting as moderator.

2. The criteria and standards to be applied are as follows:

	Criterion	Standard
a.	Comprehensiveness of the Program Definition	Taxonomy of Program Dimensions
b.	Face validity (internal consistency) of the Program Definition	Consultant employed by evaluation staff
c.	Compatibility of defined program with program environment	Values implicit in opinions of field staff

3. The Program Definition and Taxonomy of Program Dimensions are supplied to panel members in advance of the panel meeting. Field observations are reported by the evaluator at the meeting of the panel.

4. In the event that the panel does not agree in making any one judgment, the following rules are to be applied by the moderator:

   a. The evaluator must be responsible for judgments based on criterion "a" above.

   b. The consultant and the resource person from the Office of Research must be responsible for judgments based on criterion "b" above.

   c. The program director and evaluator must be responsible for judgments based on criterion "c" above.

5. A detailed record of the judgments of the panel is to be kept by the evaluator and used as the basis for a Stage I evaluation report to the program staff.

Figure V
Guidelines for Stage I Judgments

Program Definitions are weighed with the three following criteria in mind: (1) comprehensiveness, (2) face validity, and (3) compatibility.

Comprehensiveness

The basic questions here are (1) "Is there specific program information for each dimension in the taxonomy?" (Gaps should be indicated.); and (2) "Are program dimensions spelled out in acceptable form, i.e., are objectives stated in terms which describe behavior, conditions, and standards or criteria?"

Face Validity

Here the internal consistency of the definition is examined through consideration of the following questions:

1. Do the terminal objectives tell what the student does to signify success in the program?
2. Are the terminal objectives clearly related to day-to-day activities of the program?
3. Are the enabling objectives clearly related to the terminal objectives?
4. Are the entering behaviors consistent with the selection criteria?
5. Are the entering behaviors linked to the program's objectives?
6. Are the media clearly related to activities? Are they sufficient for the essential activities?
7. Are the activities related to the objectives? Is there at least one activity for each objective?
8. Are time resources realistically related to the objectives? Is there sufficient time for each specified activity?
9. Are staff qualifications and characteristics adequate for the functions defined? If not, are in-service activities spelled out to make them adequate?
10. Are all functions necessary to serve the objectives included?
11. Are the functions for each group clearly related to program objectives?

Figure VI
Criteria for Stage 1 Evaluation Judgments

12. Are staff duties clearly related to staff functions?

13. Are staff duties clearly defined?

14. Are there duties for each function?

15. Are intra-staff activities adequate for support of program operation or program objectives?

16. Are communications channels outside the program related to support needs?

Compatibility

In measuring the program's compatibility, the question to be answered is: "Is this program consistent with other programs and with the entire school system?" The basic Stage I interview questions can help determine the external consistency of the program.

1. Is sufficient time available for student participation in the program? What activity does the student give up in order to participate in the program? Does this reallocation of student time result in sacrifice of other objectives of the school program? Does it have an effect on the operation and/or goal attainment of this program?

2. Is sufficient time available for participation by the program staff and cooperating personnel? What activities do staff or cooperating nonprogram personnel sacrifice in order to participate in the program? Does this reallocation of their time result in a sacrifice of other objectives of the school program? How does it affect this program?

3. Are facilities and media now available to the program? If not, have plans been made to provide them? Is allocation of facilities and/or media to this program resulting in sacrifice of other objectives of the school program? Is this program affected by the manner in which facilities and media are allocated?

4. Are the gains for students anticipated by this program equal to, less than, or greater than possible sacrifices in other educational objectives of the school program?

5. In general, what is the effect of this program on the environment in which it operates? What is the effect of the environment upon the program?

Figure VI
Criteria for Stage I Evaluation Judgments (contd.)

D. The Process

1. Planning and arranging the meeting is the responsibility of the program evaluator. Prior to the meeting, he completes the following plans and preparations:

   a. Arrangement of details of date, time, and location of meeting and selection and inviting of participants. In choosing the consultant for the panel meeting, the evaluator confers with the Coordinator of Evaluation. Three considerations influence the choice of the consultant:

      (1) Preparation in the discipline or field of study related to the program's major emphasis

      (2) Personal capacity for a productive relationship with program leaders

      (3) Opportunity for a continuing relationship with program leaders, as affected by factors such as permanence of professional or vocational ties

   b. Preparation of participants before the meeting. At least one day in advance, the evaluator distributes the materials of the meeting and explains its purposes and procedures and the roles of participants to the consultant and to the program administrator. In communicating with the program administrator, he stresses the sharing of responsibility for evaluation judgments.

2. The meeting is opened by the status leader from the evaluation staff, who keynotes the discussion as shared responsiblilty on the part of evaluation staff, program staff, and the consultant for evaluative analysis of the program definition.

3. The status leader acts as moderator for the panel proceedings.

    a. He raises questions relative to the comprehensiveness of the program definition (see Figure VI). In the event of disagreement among the members of the panel on any one question, the moderator invokes the judgment of the program evaluator.

    b. He raises questions relative to the face validity of the program definition (see Figure VI). In the event of disagreement among the members of the panel on any one question, the moderator invokes the combined judgment of the consultant and the resource person from the evaluation staff.

    c. He calls upon the program evaluator to present the results of the Stage I field interviews. He asks for the panel's judgment of program compatibility. In the event of disagreement among the members of the panel, the moderator invokes the combined judgment of the program administrator and the program evaluator.

4. The program evaluator uses the tape recording of panel

proceedings to prepare a comprehensive outline of panel findings, relating them, item by item, to specific parts of the program definition. These findings are the culmination of Stage I evaluation and form the basis for the Stage I Evaluation Report.

5. The evaluator prepares a report of Stage I evaluation which is then distributed by the Office of Research to all members of the program staff.

Feedback: Rationale

The function of evaluation in program development and program assessment under the Pittsburgh Model is to provide both judgmental and objective information about the program for use by program decision makers. It is important that this feedback be given in a manner which encourages both the acceptance of evaluation and the use of evaluation findings for program development. Some of the guidelines which shape the form, focus, timing, and frequency of feedback are as follows:

1. Feedback is always given in a context which provides interpretation of evaluation functions and activities.
2. Feedback is given as promptly after each act of evaluation (each set of observations or judgments) as is consistent with care and accuracy of data handling. It is given in oral form

while written reports are in preparation.

3. The courtesy of a preview of each written report, before it is issued, is extended to the program administrator.
4. For every program, written evaluation reports are issued to all individuals identified as program staff or as resource personnel essential to the program.
5. In written reports, evaluation findings are presented as judgments with implications for action.
6. Interview data is returned to respondents in written form, for preview as well as for verification, before it is used for evaluation purposes.
7. Relevant feedback is given to all adult individuals within the system, whether or not on program staff, who have contributed to evaluation as subjects of observation or participants in evaluation activity.

Feedback: Description

Informal. The evaluator interacts almost continuously with program staff, both administrative and field personnel, the degree of interaction being determined both by the size and scope of the field to be covered and by the number of scheduled activities. Program activities such as in-service training meetings and group planning sessions, as well as the evaluation activities described in earlier parts of this paper, provide opportunity for informal contacts. Between scheduled activities, the evaluator visits the field operations making personal contacts with individual members of the field staff.

During these visits the evaluator seizes every opportunity for communicating recent evaluation findings and observes and records the reactions of program staff. He is systematic in directing feedback to those individuals who have given time and effort for evaluation. His contacts are most frequent with the program administrator to whom he reports not only the results of evaluation, but also the reactions of field personnel.

Formal. Feedback of this type is provided by the documents described below:

1. Evaluation Reports--At the end of each cycle of evaluation for each program, a written report is issued, with evaluation judgments and any implications for program change activity made clear. The report includes the data on which judgments are based. One section is devoted to an explanation of the nature and purpose of evaluation for the cycle. Since evaluation cycles are not tied to a specific time schedule, evaluation reports may be written as often as several times during a given school year.

2. The Annual Report--The annual report for each program satisfies the requirements of the state and federal agencies responsible

under the Elementary and Secondary Education Act of 1965 for the evaluation of Title I projects. In response to the needs and wishes of these agencies the report includes program assessment, whether or not such assessment is consistent with the evaluation model. (In anticipation of the annual evaluation report, product data is collected concurrently with process data during the school year.) In addition, the report includes a review of all evaluation activity and findings of the previous year, thus providing a view of the individual cycle evaluations as links in the spiral of program development.

3. Monthly Newsletter--As an additional means of insuring continuous feedback to all program staffs, the evaluation staff plans to institute a monthly evaluation newsletter for all personnel involved in Title I programs. The newsletter will describe (1) plans for evaluation as they evolve for each program; (2) current evaluation activities such as program definition or redefinition meetings, panel meetings, instrument development work, data collection, etc; and (3) evaluation findings. One anticipated benefit of the newsletter is the enriching of the concept of evaluation held by field personnel.

The Problems Inventory: Rationale

In Stage II, evaluation looks for lack of congruence between the defined program and the operating program. In thus leading the way to improved communications within the program staff and/or to program redefinition, evaluation functions for program development. However, a hazard to efficiency in stimulating program change activity is present in the number and range of dimensions for each program and in the current scarcity of instruments for measuring antecedent and process dimensions. For a given cycle of evaluation, thoroughness in seeking out incongruence for every program dimension can be achieved only at the cost of delay in pointing the direction of program improvement. To maximize the benefits of rapid program recycling, evaluation strategy for Stage II relies on selectiveness which is supported by a means for cultivating sensitivity to areas of probable incongruence. The problems inventory is a set of cues which facilitates the assigning of priorities for Stage II observations.

The Problems Inventory: Description

The problems inventory is developed at any convenient gathering of the program staff, or a representative sample of the staff. (Usually time is set aside for developing the problems inventory near the end of a program definition meeting). The inventory is made up of staff responses to a question such as "What single problem in the conduct of your program is most threatening to the objectives of the program?" When conditions permit, small group discussions are used to evolve a ranking for the items on the inventory.

Other Procedures

Project activity currently is aimed toward developing procedures to implement evaluation in Stage II and subsequent stages and, in addition, to facilitate evaluation staff--program staff communication and collaboration across all stages of evaluation. Specifically, procedures are being designed for:

(1) support of program staff activity for refining statements of objectives and the internal structure of program design;

(2) using field personnel for instrument development;

(3) imparting to program personnel a deeper understanding of the evaluation function and of the value to both evaluation staff and program staff of frequent interaction and open channels of communication.

EVALUATION UNDER THE PITTSBURGH MODEL: THE PREPRIMARY PROGRAM

Introduction

The Pittsburgh Public Schools enroll approximately 75,000 students and employ 1700 teachers in 108 school buildings. In 1965, 72 of these schools were designated as compensatory schools for the purpose of allocating Title I funds.

The Pittsburgh Title I effort was made up initially of 31 discrete projects. Prior to the 1966-1967 school year, the first complete year of operation, the number of projects was reduced to 28. These projects, directed toward the needs and problems of disadvantaged children, cover a wide range of strategies, including curriculum development, remediation and enrichment, innovative administrative structures, physical and mental health services, and guidance services. For the purpose of illustrating the evaluation strategy and procedures being developed in Pittsburgh, one project, the Preprimary Program is presented herein.

History

Pittsburgh's Preprimary Program was originally conceived as a crucial first step in the plan for a city-wide program of compensatory education for children from culturally and economically deprived neighborhoods. The preschool endeavor was initiated in one school in

October 1964 and extended to 12 other schools during its first six months of operation. Since March 1965, 41 additional classes have been added to the program. As of June 1967, there were 54 operating units attached to 44 elementary schools with a total enrollment of nearly 2000 children.

Description

Broadly and very loosely defined, the Preprimary Program in the Pittsburgh Public Schools seeks to provide three- and four-year olds from economically and culturally deprived areas with an opportunity to begin to develop the kinds of socio-emotional characteristics and cognitive skills they will need to succeed in school. An effort is also being made to promote physical well-being by providing the children in the program with nutritious meals and access to basic health services. During the 1966-1967 school year, the Preprimary Program employed 54 teachers, 54 assistant teachers, 54 aides, and a number of part-time volunteers. Supervision was handled by a program coordinator and two assistant coordinators. Additional staff included an art consultant, eight eurhythmics instructors, twenty staff in training, and five unassigned teachers, one of whom acted as a substitute.

The Office of Research has been involved in the evaluation of various aspects of the program since 1965. In the fall of 1966 evaluation plans and activities relative to this program became a part of the ongoing model development effort in Pittsburgh. The evaluation events and products for this program during the 1966-1967 school year, as described below, illustrate the process and products of evaluation under the Pittsburgh Model.

Program Definition

The initial definition of the Preprimary Program is based on information elicited from teachers, principals, and supervisors at two separate meetings held in February 1967. The information itself consisted of staff members' responses to several questions put to them by members of the evaluation staff. To minimize inaccuracies, discussion leaders were asked to summarize orally the essence of each participant's remarks before writing them into the record. Volunteer recorders were also utilized to insure that most of what was said would be recorded. Insofar as the discussion leaders were able to elicit and record fairly specific responses to open-ended questions about the program, the information collected represents a complete and reasonably accurate account of the proceedings of both meetings.

The first "definition meeting" took place on Friday, February 10, from 8:30 to 11:30 a.m. Twenty-four teachers were divided into six small groups to facilitate consideration of an agenda (see page 201 in Appendix A) developed by the evaluation staff. Each group consisted of four teachers, a supervisor or an administrator from the central administrative staff, and a discussion leader from the evaluation staff. The discussions were relaxed and spirited, and each of the six working groups produced a lengthy list of fairly specific information about the program. Toward the end of the morning, the participants were given an opportunity to voice their opinions about problems confronting the program (See Appendix B, pages 215-217,for Problems Inventory.)

Another meeting to complete the definition of the program was held on Friday, February 24, from 8:30 to 11:30 a.m. Twelve teachers were divided into three groups to discuss questions about the program. An administrator, a supervisor, and a principal invited by the evaluation staff rounded out each of the small groups. It was hoped that by asking principals to participate their perspective--albeit neglected at the first meeting--might still be incorporated into the final definition of the program.

A combination of fewer participants and more observers, however, had an adverse effect on the mood of the discussion and the content of the information obtained. For the most part, the participants' responses to questions put to them by the evaluation staff were much less specific than those of their counterparts at the first meeting.

The recorded proceedings of all of the small group sessions at the two definition meetings were synthesized and written up by the program evaluator as the first formal definition of the program (see Appendix C, pages 221-234). The program definition was issued by the evaluation staff all members of the program staff on April 13, 1967.

As a check on the accuracy of the evaluation staff and a way of recording any divergent opinion not expressed during the definition meetings, an opportunity was provided, by means of a mailed questionnaire, for program staff to react to the definition meeting and to the first formal definition of the program. The instrument (open-ended questions, see Appendix D on page 237) was mailed out a month after the mailing of the definition resulting in a low rate of return.

The reactions of teachers who responded to the questionnaire were unanimously favorable to the program definition, their most frequent comments being "well done," "accurate," and "a helpful guide." Some suggested that there was too much emphasis on the economic aspects of cultural deprivation, that the definition tended toward overgeneralization, and that certain dimensions including job descriptions needed to be amplified.

In an attempt to elicit more specific comments about the definition, the program evaluator interviewed a number of staff members informally. Their criticisms and suggestions are summarized below:

> It was suggested that a major objective of the Preprimary Program is the development of greater verbal ability and, more specifically, greater fluency in the kind of language these children will need to succeed in school and in later life. Through their participation in a "rich experiential program," it is hoped that they will begin to develop the conceptual understandings as well as the vocabulary they need to make more effective use of their language.

To this end every opportunity is taken to expand the children's range of experiences and their vocabulary. Snack time, neighborhood walks, and trips are mentioned by teachers and coordinators alike as activities which are particularly appropriate for these purposes. "Show and tell" and sentence completion exercises, on the other hand, are not considered appropriate activities for three- and four-year olds.

Manipulative activities are described as providing deprived youngsters with manageable tasks which they can successfully complete, as well as fostering small muscle development and better eye-hand coordination.

The interviewees also suggested that the definition be modified to include a description of how OEO criteria are applied to select participants for the program and more detail about parent meetings, in-service training, and the newly instituted feeding program. With regard to selection, the interviewees said that children are picked from a list of eligible applicants by a panel consisting of the preprimary teacher, the school principal, the community agent, and a representative of the local community action organization. They also objected to the fact that income is all too often the sole determinant of eligibility for the program. Most of those interviewed wished to point out that cultural backwardness is not always the result of economic deprivation and that other factors should be given equal weight in the selection of program participants.

> The program staff members interviewed were also quick to point out that prior experience in traditional teacher-centered classrooms could be a disadvantage to the teacher in the child-centered Preprimary Program and that all personnel involved in the program have to be willing and able to work closely with other adults.

As a consequence of program staff reaction to the initial definition of the program, the definition was refined and extended. The revised version appears on pages 179 through 195 of Appendix A.

Evaluation Judgments

In preparation for the making of evaluation judgments appropriate to the first stage of evaluation, arrangements were made by the program evaluator for the collection of interview data relevant to program compatibility and for the conduct of a Stage I Panel Meeting.

Nineteen face-to-face interviews and one telephone interview were conducted in seven representative schools in May and June of 1967. Seven preprimary teachers, seven kindergarten teachers, four principals, and one assistant principal were asked questions about the use of facilities and media, the availability of qualified personnel, the use of staff time, and the impact of the Preprimary Program on other programs in each of the schools. Interviewees were encouraged to make judgments about the compatibility or incompatibility of the program within the total system. The interview schedule appears on page 203 in Appendix A. The responses of the interviewees are on pages 205 through 212 in Appendix A. Briefly, the opinion of the interviewees was as follows:

> With the possible exception of equipment for the development of large muscles, facilities, equipment, materials, and supplies are generally adequate to meet the needs of the program. With respect to access to auxiliary services, these are also generally available when needed. The recent decision by the Allegheny County Health Department to provide health services to all preschoolers meets a need long felt by all teachers interviewed. Most of the respondents indicated that they need more time to plan and coordinate their activities. The attitudes of principals and other teachers about the Preprimary Program and, as a result, the extent to which the program fits into the overall school operation, seem to vary as their understanding of its purposes, methods, and special problems.

A Stage I Panel Meeting was held for the Preprimary Program on Thursday, May 4. Members of the panel were the Director of Compensatory Education (the chief administrator of the Preprimary Program), a consultant in preschool education employed by the evaluation staff, and the following representatives of the evaluation staff: (1) the program evaluator, (2) a research assistant skilled in writing educational objectives and in program design, and (3) the Coordinator of Evaluation.

Findings of the panel relative to the criteria of comprehensiveness and face validity are presented on pages 179 through 195 in Appendix A in a format which relates them, item by item, to specific parts of the program definition. In summary, the findings were as follows:

> Comprehensiveness. In order for the Preprimary Program definition to serve as a useful basis for further evaluation of the program, it must be made more comprehensive. For example, objectives are not sufficiently specific and are not stated in behavioral terms. No attempt has been made to distinguish between terminal, ultimate, and enabling objectives.
>
> Once more specific objectives have been established, the antecedent and process dimensions of the definition can be amplified and their relationship to program objectives spelled out in considerable detail. For example, the entering behaviors do not presently describe those characteristics of children which will be modified as a result of their participation in the program and the kinds of deficiencies the program is designed to overcome. The definition is not as specific as it might be in describing other antecedent dimensions including staff qualifications, administrative support, facilities, media, and time constraints and how each of these affects the achievement of program objectives.
>
> Finally, the process dimensions of the definition, the day-to-day activities of the children and the program staff through which human and material resources are combined to produce specified outcomes, are neither complete nor closely related to outcomes and antecedents. The enabling and terminal objectives to which each activity contributes have not been identified.

> Validity. The Preprimary Program definition as it stands is difficult to assess in terms of internal consistency. Once it has been rewritten in the manner described above, it should be easier to judge the face validity or the reasonableness of the functional relationships set forth.

In applying the criterion of compatibility, the panel reviewed the interview data reported above and concluded that no incompatibilities between the Preprimary Program and others in the system had been identified.

At the conclusion of the Preprimary Panel Meeting, the evaluation staff assessed the effectiveness of the meeting in the following manner:

1. The judgment process was impeded by the nature of the materials being used. The Preprimary definition presented to the panel was in its original, narrative form and thus it was difficult for the panel to identify, examine, and relate the specific dimensions.[7]

2. In terms of the credibility of panel findings, it appeared to the evaluation staff that the program administrator accepted both the findings and the need for action based on their findings. However, the evaluation staff was aware of her uncertainty as to how change activity would occur.

---

[7] As a result of this experience, all program definitions subsequently issued by the evaluation staff were prepared in outline format.

Reporting

The action of the panel in making the Stage I evaluation judgments brought to an end the first cycle of evaluation for the Preprimary Program. An evaluation report was prepared by the program evaluator and distributed by the evaluation staff to all members of the preprimary staff. This report accomplished the following purposes:

1. To explicate and interpret Stage I evaluation to the preprimary staff
2. To present all instruments and data: (1) the Taxonomy of Program Dimensions, (2) the interview schedule for the preprimary definition meeting, (3) the specific dimensions of the Preprimary Program, (4) the panel findings relative to specific dimensions of the Preprimary Program, (5) the interview schedule for the Stage I field interviews, and (6) the interview responses of field staff
3. To present evaluation findings and judgments as implications for change activity

A copy of the report appears as Appendix A.

Stage II Planning

The first cycle of evaluation of the Preprimary Program came to an end, coincidentally, near the end of the school year. However, enough time remained before the ending of the session for beginning the development of a Stage II instrument for use in the next cycle of evaluation.

Rationale for Stage II Instrument

Stage II evaluation attempts to ferret out incongruencies between the current program definition and the operating program in their antecedent and process dimensions. One area of possible incongruence was signaled by Item 8 of the Problems Inventory, produced by the preprimary staff at the first definition meeting in February 1967 (see Appendix B, pages 215-217). Item 8 reads in part as follows:

> There is some concern [among the preprimary staff] about pressures toward greater emphasis on structured learning of a cognitive nature.... This was reflected in differences of opinion regarding the usefulness of detailed planning and record-keeping for each individual child. All participants are willing to be held accountable for what they are doing; most see planning, keeping logs, and standard tests as useful aids which help them become more effective teachers. More specifically, logs were defended [by supervisors and some teachers] as means of making the teacher's objectives and methods explicit, thereby helping to insure consistency and continuity in the Preprimary Program. They also serve as a place to record the kinds of progress each child is making as well as the areas in which he may need extra attention.

The lack of agreement among the preprimary staff which is evident in Item 8 suggests a lack of uniformity and/or consistency in implementing the program, particularly in respect to the methods used by teachers. Since teacher methods might be inferred from the kinds of activities in which children participate and in the kinds of interaction that take place between children and adults, it was decided to observe these aspects of the operating program.

Development of Observation Categories

In order to insure that those who observe in the field attend to relevant and only relevant behaviors and that all observers record similar behaviors in the same way, it is necessary to set up categories of children's activities and adult-child interactions. A prior stage is making preliminary detailed observations of classroom activity to determine the range of activities and interactions that must be subsumed by the categories.

Preliminary observations were planned and implemented in June 1967. Each of four observers trained in the observation of preschool children at the Arsenal Child Study Center visited two classrooms in two different schools, one in the morning and one in the afternoon, for a period of four days. Each observer watched two children at play for two twenty-minute intervals at every session he attended.

To insure a measure of uniformity in the content of the observations made and recorded, as well as to alert the observers to the kinds of information considered important by the evaluation staff, a list of guidelines was drawn up and distributed to each observer. These guidelines are presented in Appendix E, pages 241-242.

The eight schools at which observations were conducted were selected by using a table of random numbers. Principals were informed by letter and by telephone of the purposes of the observations and the methods to be used.

The observations were written up in the form of a chronological narrative by the observers and submitted to the evaluation staff for detailed analysis. These reports are presently being utilized to formulate a list of specific categories of play activities and teacher-child interactions which will serve as a basis for more intensive observations to be carried out during the next cycle of evaluation.

Program Assessment

During the summer following the end of the first cycle of evaluation for the Preprimary Program, the annual evaluation report to state and federal agencies was prepared. In anticipation of the requirements of the annual report and of the need for instruments for Stage III evaluation, product data for the Preprimary Program had been collected during the 1966-1967 school year. Development of the instruments had begun even before the fall of 1966.

The Socio-Emotional Scales

A major purpose of the Preprimary Program since its inception has been the development of the socio-emotional characteristics children need to succeed in school. From January 1966 to April 1966, 12 preprimary teachers met one afternoon each week with a representative of the evaluation staff to formulate specific behavioral characteristics of a socio-emotional nature that children can reasonably be expected to acquire as a result of their participation in the program and to develop an instrument with which to measure their progress over the course of a school year. More specific behavioral objectives were set forth and four sequential levels of achievement established for each objective.

The scales were empirically tested in preschool classrooms to determine proper sequencing and develop the most appropriate terminology with which to describe student behavior at each level of achievement. An attempt to further refine the scales by describing as many objectives as possible in terms of the six major curriculum based activity areas was also made before they became available for use by the program staff. This improvement makes it possible to use the scales to determine the kinds of progress each child is making toward four of the nine objectives in all six activity areas. A complete description of the socio-emotional scales and directions for their use are contained in Appendix F, pages 245-248.

Results of First Administration. During the 1966-1967 school year, eight teachers recorded their observations of students' progress in the direction of socio-emotional adjustment on two or more separate occasions. Table 1 shows the means and standard deviations for each of the eight schools on each scale as well as the means and standard deviations for the total group both times the scales were administered.

TABLE 1

Class Mean and Standard Deviation For the Social-Emotional Scales 1966-1967

	Maximum Score		Total Group		Arlington		Columbus	
Number of Students			200		26		11	
Administration			1	2	1	2	1	2
Activity Area								
Large Muscle	12	Mean	7.14	9.28	8.88	10.54	4.91	7.00
		Std. Dev.	2.96	2.55	2.38	1.65	1.86	2.32
Creative	12	Mean	6.79	8.85	7.00	9.35	6.27	10.00
		Std.Dev.	2.73	2.64	3.28	2.24	2.37	1.39
Dramatic Play	12	Mean	7.03	9.34	7.15	9.31	6.64	10.09
		Std. Dev.	3.12	2.87	3.41	3.47	2.20	1.81
Manipulative	12	Mean	7.23	9.24	8.23	9.54	5.09	7.73
		Std. Dev.	2.76	2.45	3.02	2.56	2.02	2.57
Sensory	12	Mean	7.15	9.28	8.26	10.12	6.45	9.91
		Std. Dev.	2.75	2.59	2.44	2.77	2.16	1.92
Cognitive	12	Mean	6.50	8.36	8.31	10.35	6.18	9.45
		Std. Dev.	3.39	3.25	3.41	2.57	2.18	2.01
Behavior Scale								
Initiates own activities	18	Mean	11.02	13.64	13.35	15.00	5.55	1.73
		Std. Dev.	4.50	3.96	3.27	3.97	4.03	3.92
Relates to Group	18	Mean	11.12	14.09	10.50	13.30	11.55	5.82
		Std.Dev.	4.28	3.82	4.73	4.26	2.42	1.66
Stays with a task	18	Mean	10.03	13.28	11.96	14.57	9.00	3.72
		Std. Dev.	4.09	4.14	5.15	4.38	2.76	3.55
Awareness of Changes	18	Mean	9.75	13.16	11.88	14.73	9.09	3.36
		Std. Dev.	5.10	4.72	6.52	5.40	4.90	3.35
School Environment	18	Mean	16.47	17.43	16.62	17.77	14.73	16.91
		Std. Dev.	3.06	2.21	3.08	1.17	3.13	2.42
Shares Objects and Affections	18	Mean	11.25	13.47	9.00	12.00	14.18	16.36
		Std. Dev.	4.16	3.73	3.05	3.79	3.02	2.80
Controls Emotions	18	Mean	8.19	11.49	7.15	10.38	10.36	15.82
		Std. Dev.	6.54	5.85	7.20	7.11	6.62	3.03
Adapts to Changes	18	Mean	15.92	16.08	15.92	16.85	11.4	14.18
		Std. Dev.	4.50	3.84	3.77	3.40	6.2	6.16
Takes Care of Personal Needs	18	Mean	11.19	13.02	14.76	16.15	12.55	14.18
		Std.Dev.	4.15	4.04	4.23	2.82	4.20	4.04
Totals	162	Mean	9.73	12.00	10.60	12.66	8.93	12.42
		Std.Dev.	4.96	4.54	5.18	4.66	4.78	4.32

TABLE 1 (Contd.)

Class Mean and Standard Deviation
For the Social-Emotional Scales 1966-1967

Crescent		Forbes		Fort Pitt		Frick		Miller		Sheraden	
24		31		34		36		20		18	
1	2	1	2	1	2	1	2	1	2	1	2
7.17	8.92	6.65	9.16	7.12	8.50	7.14	9.97	6.95	9.20	7.00	8.72
2.14	2.15	2.75	2.20	2.78	3.03	2.98	2.09	3.83	3.36	3.78	4.03
7.54	7.58	6.00	8.35	7.35	8.74	6.11	8.78	7.00	8.15	7.39	10.17
1.61	1.99	2.80	3.37	2.57	2.91	2.33	2.48	3.58	3.40	2.85	2.72
8.13	9.67	5.90	8.90	8.00	9.09	6.94	9.78	5.70	8.70	7.39	9.61
2.52	2.46	2.93	2.89	2.55	2.75	3.46	2.28	3.65	3.77	3.08	3.18
8.37	9.25	6.19	9.00	7.65	9.03	6.77	9.88	6.60	8.85	8.28	10.33
1.91	1.59	2.89	2.44	2.39	2.72	2.55	1.78	3.50	3.13	2.53	2.67
7.50	7.92	6.71	9.32	6.64	8.62	6.78	9.89	6.50	8.85	8.17	10.00
2.02	1.79	3.19	2.52	2.81	2.82	2.55	1.76	3.48	3.42	2.99	2.78
7.46	8.04	6.52	9.29	3.44	5.44	7.00	8.22	5.95	7.55	8.22	10.38
3.02	2.99	3.00	2.38	2.72	3.33	2.59	2.34	4.50	3.88	2.57	3.01
11.25	12.83	8.97	12.65	13.18	14.44	10.50	13.94	10.10	12.80	12.28	14.39
3.55	2.61	4.84	5.12	3.78	3.94	3.37	2.60	5.31	4.40	4.76	4.40
12.63	14.29	10.80	14.25	11.35	13.88	9.97	13.75	10.80	13.10	12.50	15.78
2.37	2.03	4.77	3.61	4.13	4.47	3.76	3.51	5.97	5.15	4.34	3.50
11.58	12.29	9.61	13.06	9.24	11.73	9.97	14.16	8.15	11.65	10.22	15.83
2.44	3.29	4.49	4.17	3.60	4.48	3.44	2.43	4.55	5.72	4.55	3.39
11.04	12.17	8.42	14.51	6.73	9.52	10.17	14.06	9.75	13.90	12.56	14.00
2.72	2.11	5.13	3.52	4.20	5.52	4.24	3.64	5.21	4.54	5.52	5.98
17.00	18.00	16.26	17.03	17.29	17.47	16.33	18.00	16.50	17.40	15.66	16.00
2.88	0.00	3.53	3.13	1.96	2.27	2.72	0.00	3.30	1.84	4.18	4.11
11.75	12.75	10.25	13.16	12.88	14.12	11.67	14.67	10.50	12.60	10.67	12.67
4.50	4.07	3.85	3.25	3.00	2.91	4.28	3.34	5.46	4.72	4.39	4.05
10.50	11.00	5.61	9.48	10.59	12.88	8.33	13.50	6.00	8.70	7.33	11.00
7.13	6.29	5.35	5.53	6.10	4.21	6.13	4.15	7.28	8.13	5.26	5.14
14.00	14.25	16.65	17.03	17.65	17.65	13.17	16.00	11.70	13.50	17.00	17.00
5.50	4.61	2.98	2.24	1.43	1.43	3.74	3.20	5.99	5.79	3.08	3.08
7.25	7.75	11.80	12.97	8.12	10.94	11.50	13.67	11.10	13.20	14.67	17.33
2.48	3.74	2.44	2.72	3.87	3.75	2.21	2.72	2.19	2.46	5.13	1.44
10.21	11.11	9.09	11.88	9.82	11.47	9.49	12.53	8.88	11.21	10.62	12.88
4.41	4.25	5.09	4.45	5.11	4.86	4.44	3.91	5.43	5.19	5.09	4.66

Reported Conclusions. The annual report presented the following conclusions:

> It is presently impossible to draw conclusions from the data as to student performance even though the class means on all scales are uniformly higher the second administration than the first. However, the subjective nature of the data must not be ignored. There is some evidence that the scales have been improperly applied and that some teachers are producing unreliable ratings. It therefore becomes necessary to engage in further teacher training in the administration of the scales and to secure evidence as to the reliability of new ratings.

The Preschool Cognitive Test

Since language skills and conceptual development are critical factors in school success and are areas in which children from deprived neighborhoods show glaring deficiencies, another major objective of the Preprimary Program is to improve the ability of individual children to express, communicate, and understand. In the absence of agreement among staff members as to the importance and place of more specific cognitive objectives, however, it was left to the evaluation staff to develop a list of specific objectives and a diagnostic instrument with which to assess individual progress in meeting these objectives.

After a review of the literature on the subject and a careful examination of the reading curricula to which children are ordinarily exposed in the first grade, a member of the evaluation staff identified a number of cognitive skills that most school systems seem to assume children will have begun to master before they begin their formal education. It was decided to make these skills the basis of a diagnostic cognitive test. By changing the items found in standard reading readiness tests from group paper and pencil exercises to individual pointing or verbal responses, the staff was able to develop a reading diagnostic test that could be administered to four-year olds.

Test items were created for ten identifiable skills and pretested on a group of 60 non-disadvantaged private nursery school children. Items that were passed by at least 75 percent of the nursery school children and at least 40 percent of the deprived group were included in a printed version of the test that was administered on a somewhat larger scale in October and November of 1966.

This version of the test contained 130 items designed to test preschoolers' knowledge in seven basic skills areas. These areas included vocabulary, classification and generalization, visual and auditory discrimination, sequences, basic numerical concepts, and directional - spatial relationships. (For an outline description of the 1966 Preschool Cognitive Test see Appendix G, pages 251-255).

Results of First Administration. Test results were obtained from 428 deprived four-year olds, 326 non-deprived four-year olds, 55 deprived five-year olds, and 60 non-deprived five-year olds. Table 2 presents some of the findings which have relevance for test development.

TABLE 2

Number of Items Passed by at Least 40 Percent of the Group (N=130)

Group	No. of Items Passed
Deprived Preschool	87
Non-Deprived Preschool	121
Deprived Kindergarten	107
Non-Deprived Kindergarten	127

The test results have been used by the evaluation staff to eliminate those items on the test which do not discriminate between deprived and non-deprived students or between preschoolers and kindergarten children and those items that show no variability (i.e., those which everyone passed or everyone failed). Preparations are being made to administer a modified version of the test to a sample of comparable size during the coming school year. The new version contains 113 of the original 130 items.

Reported Conclusions. The annual report included the following statements relative to conclusions:

> ...plans are underway to administer a modified version of the 1966 test to a sample of comparable size sometime in the fall of 1967. During the period of test development, no definite conclusions concerning program effectiveness have been attempted.

The Peabody Picture Vocabulary Test

Efforts to evaluate the effects of the Preprimary Program on children from deprived neighborhoods began in 1965 when the Office of Research administered a battery of psychological tests to 203 three- to six-year olds in seven schools. The tests administered were the Peabody Picture Vocabulary Test (PPVT), the Children's Apperception Test, the Goodenough Draw-A-Man Test, and an experimental check list. On the basis of the initial findings, which were reported in September 1965, it was decided to test the same children again a year later to gauge the impact of participation in the program on their intellectual ability. Because its relationship to later school achievement is more firmly established than those of any of the other tests mentioned, the Peabody Picture Vocabulary Test was selected for this purpose.

The testing program was initiated in 1964-1965 with a group of children who showed below average verbal functioning (mean I. Q. 85.2). In 1966, after 12 to 20 months in the program, a 59 percent sample of the 203 children originally tested showed an average I. Q. of 97.0. However, despite the significant improvement on the part of many children, those who had higher than average I. Q.'s on the first testing either remained the same or scored lower on the retest.

A number of possible explanations were sought to account for the performance of high I. Q. children. To the extent that students who performed well on the test were operating closer to their real potential than those who did not, it would be reasonable to expect the high I. Q. children to either remain the same or improve less than the lower I. Q. children. However, this explanation does not account for the significant decrease in I. Q.; regression effects remain as a possible explanation.

In examining the procedure used to determine I. Q. from the raw score of the PPVT, it was discovered that the I. Q. conversion tables were relatively crude. If one was not careful to administer the test an equal number of months apart to all members of a sample, comparability within the sample would be invalid. To circumvent this problem, the evaluation staff interpolated the I. Q. conversion tables supplied with the test to make them more sensitive to time differences. The 1965 and 1966 test results were then rescored using the interpolated tables. It was found that the tendency of the high I. Q. children to score lower disappeared.

Results of Retest Using Interpolated Tables. Using the interpolated tables, the 1965 data yielded an average I. Q. of 86.32. One year later, the mean I. Q. of students who had been in the program from 12 to 20 months was 99.57. The mean I. Q. change was 13.25 I. Q. points.

Results of Second Retest. To determine the stability of the improvement in I. Q., the PPVT was administered again in May and June of 1967. The general improvement in I. Q. first seen after one year in the program was maintained. The mean I. Q. in 1967 was 100.50. Mean scores and changes for each school included in the sample are shown in Table 3.

Table 4 shows the distribution of scores for the entire I. Q. range. The table indicates that a larger percentage of students fall in the 90-109 range after exposure to the Preprimary Program. The number of students in the low I. Q. ranges also decreased markedly (from 18 percent in 1965 to 2 percent in 1967). With regard to students with I. Q.'s over 100, Table 4 indicates the change for these students was small compared to the increase for the total group.

TABLE 3

Mean I.Q. Score and Mean I.Q. Change by School (Interpolated Scores)

	1967 Testing				1966 Testing			1967 Testing		2nd-3rd Testing MEAN IQ CHANGE	1st-3rd Testing MEAN IQ CHANGE
	Total Group (Uninterpolated)		Sample Group (Interpolated)		Sample Group (Interpolated)		Testing MEAN IQ CHANGE	Sample Group (Interpolated)			
	N	MEAN	N	MEAN	N	MEAN		N	MEAN		
Forbes	28	95.25	10	96.90	10	105.60	+ 8.70	10	111.00	5.40	14.10
Letsche	28	77.57	15	79.06	15	94.86	+15.80	15	94.13	-.73	15.07
Madison	32	86.84	20	90.45	20	102.90	+12.45	20	101.00	1.90	14.16
McKelvey	33	83.97	16	86.00	16	100.25	+14.25	16	101.12	.87	15.12
Miller	26	83.31	10	85.10	10	101.50	46.40	10	99.00	-2.50	13.90
Vann	25	82.68	7	82.28	7	92.14	+9.86	7	102.14	10.00	19.86
Weil	37	85.94	18	84.44	18	97.67	+13.23	18	99.05	1.38	14.61
Total	203	85.15	96	86.32	96	99.57	+13.25	96	100.50	+ .93	+14.18

TABLE 4

Number and Percentage of Children in Re-Test Sample Falling in Conventional I. Q. Categories on First, Second, and Third Testing (Interpolated Scores)

I. Q. Category	1965		1966		1967	
	N	%	N	%	N	%
120 and above	2	2	7	7	9	9
110-119	7	7	20	21	13	14
100-109	15	16))37%	27	28))45%	30	31))57%
90-99	20	21)	16	17)	25	26)
80-89	14	15	17	18	11	11
70-79	21	22	6	6	6	6
69 and below	17	18	3	3	2	2

Reported Conclusions. The annual report presented the following conclusions:

> Results from Peabody Picture Vocabulary Tests administered in 1965, 1966, and 1967 indicate a marked improvement in I. Q. on the part of children who have been exposed to the Preprimary Program.... With regard to students with initial I. Q.'s over 100, the findings suggest that either these students were operating near capacity initially, or the program did not address itself to them. Other explanations may be forthcoming. The Office of Research will continue to monitor the performance of selected students during the 1967-1968 school year to find whether their achievement measures up to the potential shown by I. Q. scores. Plans are also being made to collect data from children who are not in the program for control purposes. It may be possible to determine the extent to which marked improvement in I. Q. scores can be attributed solely to participation in the Preprimary Program.

Planning for Cycle II

When preprimary sessions reconvene in September 1967, the second cycle of evaluation for the program will begin. In the meantime, the rationale and detailed plans for this cycle are being prepared by the program evaluator with assistance from the evaluation staff. The preliminary plan for Cycle II evaluation is presented in chart form in Appendix H, pages 258-259. The specific evaluation activities implied by the cycle plan are shown below in Table 5.

TABLE 5

Preprimary Evaluation
Time Line, Cycle II
1967-1968

Evaluation Activity	Begin	End
Arrange for and conduct modified definition meeting	Oct. 2	Oct. 18
Draw sample for fall administration of Preschool Cognitive Test	Oct. 2	Oct. 18
Code and process data collected from preliminary classroom observations	Oct. 2	Oct. 20
Analyze and synthesize products from modified program definition meeting: prepare revised program definition	Oct. 19	Nov. 1
Develop observation checklist for preliminary data	Oct. 23	Nov. 1
Identify and train classroom observers to use observation checklist	Oct. 23	Nov. 10
Mail revised definition and response form to program staff	Nov. 3	
Develop interview schedule concerning support and communications dimensions of the operating program	Nov. 3	Nov. 10
Make informal contacts with program staff to gauge reaction to revised definition	Nov. 13	Nov. 24
Conduct interviews	Nov. 13	Dec. 1
Conduct classroom observations	Nov. 13	Nov. 22
Identify and train raters to administer socio-emotional scales; draw sample of students	Nov. 13	Nov. 17
Administer Preschool Cognitive Test	Nov.	Nov.
Modify revised definition to incorporate comments and criticisms	Nov. 24	Dec. 1
Arrange for and conduct panel meeting to judge revised definition	Nov. 24	Dec. 6
Administration of socio-emotional scales by teachers and other raters	Late Nov.	Early Dec.
Code, process, and analyze data from classroom observations to determine congruence between definition and operating program	Nov. 27	Dec. 15
Analyze interview data to determine congruence between definition and operating program	Dec. 1	Dec. 15
Prepare and mail report of panel proceeding to program staff	Dec. 6	Dec. 15
Collect socio-emotional data after fall ratings; analyze to determine reliability of scales and to describe entering behavior of students	Dec. 15	Jan. 10
Complete necessary modification of scales in preparation for winter administration	Jan. 10	Jan. 24
Prepare and mail Cycle II Evaluation Report	Jan. 10	Jan. 24
Plan Cycle III evaluation	Jan. 10	Jan. 24

APPENDIX A

EVALUATION REPORT, STAGE I

PREPRIMARY PROGRAM

JOHN V. O'NEILL, EVALUATOR

I. Introduction: Purpose of this Report

Whereas provisions of the Elementary and Secondary Education Act require annual evaluations of all Title I programs and reports to governmental agencies, local decision-making needs are best served by continuous evaluation and more frequent reporting. The Office of Research is conducting an ongoing evaluation of each Title I program, including the Preprimary Program. To serve the needs of each program staff, it is the intention of this office to issue a report at the conclusion of each phase of evaluation so that decision-making implications can be examined as they become apparent. The present report marks the end of the first phase in the evaluation of this program.

II. Nature and Purpose of Stage I Evaluation

The purpose of evaluation in the first phase is to judge the specifications or the "definition" of the program by applying criteria used with all Title I programs. Two kinds of observations, or data, are collected to accomplish this purpose.

The first set of observations consists of the answers of members of the program staff to questions about the program. These questions are based on an exhaustive list of program elements developed by the Office of Research as a standard to be used in evaluation (see Appendix A, page 199).

For the Preprimary Program, the necessary information was obtained from teachers, principals, the program coordinator and her assistants, and three representatives from the Department of Curriculum and Instruction at two meetings held at the Board of Education in February 1967.

The first "definition meeting" took place on Friday, February 10. Twenty-four teachers selected by the program coordinator and her assistants were divided into six small discussion groups to facilitate consideration of an agenda developed by the Office of Research. (See Appendix B, page 201). Each group consisted of four teachers, a Coordinator or a representative from the Central Office staff, and a discussion leader appointed by the Office of Research. The discussions themselves were relaxed and spirited, and each of the six working groups produced a lengthy list of fairly specific information about the program.

To complete the formal definition, 12 teachers were invited to a second meeting held on Friday, February 24. It was hoped that by asking three principals to participate, their perspective--albeit neglected at the first meeting--might still be incorporated into the program definition. A combination of fewer participants and more observers, however, had an adverse effect on the mood of the discussions and the content of the information obtained. For the most part, the participants' responses to questions put to them by representatives of the Office of Research were much less specific than those of their counterparts at the first meeting.

After these two "definition meetings," the recorded proceedings were synthesized and put into narrative form by the Office of Research.

This preliminary definition of the Preprimary Program was mailed to all members of the staff on April 13, 1967.

The second set of observations consists of the answers of program and non-program staff in the schools to questions about the manner in which the program (as defined by the program staff) fits into the overall school setting. For the Preprimary Program 19 face-to-face interviews and one telephone interview were conducted in seven representative schools in May and June of 1967. Seven preprimary teachers, seven kindergarten teachers, four principals, and one assistant principal were asked questions about the use of facilities and media, the availability of qualified personnel, the use of staff time, and the impact of the preprimary program on other programs in the school. They were encouraged to make judgments about the compatibility or incompatibility of the program within the total system.

The questions used for the interviews, the interview schedule for the Preprimary Program, appear in Appendix C, page 203. The responses of the interviewees are reported in Appendix D, pages 205-212. Very briefly, the opinion of the interviewees was that with the possible exception of equipment for the development of large muscles, facilities, equipment, materials, and supplies are generally adequate to meet the needs of the program. With respect to access to auxiliary services, it was reported that these are also generally available when needed. The recent decision by the Allegheny County Health Department to provide health services to all preschoolers meets a need long felt by all teachers interviewed. Most of the respondents indicated that they need more time to plan and coordinate their activities. The attitudes of principals and other teachers about the Preprimary Program and, as a result, the extent to which the program fits into the overall school operation seem to vary as their understanding of its purposes, methods, and special problems.

Having collected a definition of the program and the opinions of field staff regarding its compatibility with the school setting, evaluation turns toward the application of criteria for judging the program definition. For the first phase of evaluation (Stage I) the following criteria are specified:

1. Comprehensiveness of the program definition
2. Face validity of the program definition
3. Compatibility of the program, as defined, with the program environment

The procedures for applying the Stage I criteria are spelled out by the Office of Research as follows:

1. Judgments are made for each program by a panel consisting of the program evaluator, the program director, one consultant employed by the Office of Research, and one resource person from the Office of Research, with the Coordinator of Evaluation acting as moderator.

2. In applying the criterion of comprehensiveness, the standard to be used is the Taxonomy of Program Dimensions (see Appendix A, page 200). In the event that the panel does not agree, the evaluator must be responsible for judging the definition.

3. In applying the criterion of face validity, the standard of logic (reasonableness) as viewed by the consultant is to be used, except that both the consultant and the resource person from the Office of Research must agree.

4. In applying the criterion of compatibility, the panel will consider the perceptions and opinions of field staff as reported by the evaluator. The standard implicit in these opinions is the order or hierarchy of objectives for the entire school system. In the event that the panel does not agree, the program director and evaluator must be responsible for the judgment.

A Stage I Panel Meeting was held for the Preprimary Program on Thursday, May 4. The members of the panel were Mr. Vin O'Neill, program evaluator; Miss Phyllis Lewis, Director of Compensatory Euccation; Miss Judy Taylor, Research Assistant; Mrs. Mary Jane Duda, Coordinator of Evaluation; and Mrs. Lauren B. Reznick, Consultant.

Mrs. Resnick received her A. B. degree magna cum laude from Radcliffe College and her A.M. and Ed.D from Harvard's Graduate School of Education. She has done basic and applied research in the areas of cognition and motivation related to social studies instruction and learning and instructional procedures, including programmed instruction, syntactic comprehension, and the development of teacher training programs based on the concepts of behavioral control. While a staff consultant and senior scientist at Basic Systems, Inc. in New York, she prepared training materials for instructional programming and designed remedial curricula and special materials in reading for dropouts. In addition to having been a research assistant at the Harvard Graduate School of Education, Mrs. Resnick has been a teacher of social studies at the American School in Paris, a research associate at Harvard's Laboratory for Research in Instruction, and a lecturer at the City University of New York. In addition to her present role as a consultant to the Office of Research, she has held consultantships to Mobilization for Youth and the Center for Applied Linguistics. She is presently employed as a research associate at the University of Pittsburgh's Learning Research and Development Center and is the Director of Research for its experimental public school program for three- to eight-year olds. She has included in her very active career the publication of numerous articles, papers, and book reviews.

III. Findings

Findings of the panel relative to the first two criteria are presented in the following pages in a format which relates them, item by item, to specific parts of the taxonomy and of the program definition.

PROJECT PREPRIMARY

Page 1

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy GENERAL

Specific Dimensions	Program Definition	Judgments
I. Overall Statement of Objectives and Rationale for the Program	The program is designed to provide three- and four-year olds from economically deprived neighborhoods with opportunities to begin to develop the kinds of socio-emotional characteristics and cognitive skills children need to succeed in school.	The definition as it now stands is too general to be of much use as a basis for evaluation without considerable modification and expansion. Socio-emotional objectives are not spelled out in enough detail. Cognitive objectives are not set forth in behavioral terms. Antecedent and process dimensions are too sketchy and are not tied closely enough to program objectives.
II. Description of Scope		
A. Number of Pupils and Schools Involved	As of May 1967, there were 56 operating units attached to 44 elementary schools with a total enrollment of approximately 1600 children.	
B. General Description of Staff	1. Classroom personnel (for each unit) a. One teacher b. One assistant teacher c. One aide d. One (or two) volunteers 2. Supervisory and other staff a. A coordinator and two assistant coordinators b. A eurhythmics staff consisting of one supervisor and seven teachers c. One art consultant d. Twenty staff in training e. Five unassigned teachers, one of whom has been acting as a substitute	

PROJECT PREPRIMARY

Page 2

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy OUTCOMES

Specific Dimensions	Program Definition	Judgments
I. Objectives--changes expected to occur in program participants as a result of the program	At the end of the Preprimary Program, participants should do the following things: A. Show signs of greater muscle control and coordination B. Have begun to develop some of the socio-emotional characteristics normally acquired during the preschool years 1. Emotional security a. Be able to control emotions in frustrating situations b. Be able to adapt to changes in routine c. Be able to pursue activities independently d. Be aware of changes in environment e. Be able to stay with tasks until they are completed 2. Social adjustment a. Relate to a group and participate in group activities b. Show trust by sharing objects and affections c. Show trust by accepting the school environment d. Take care of personal (physical and emotional) needs in a socially acceptable way	Terminal, ultimate and enabling objectives are not differentiated. Behaviors are not specific. What does a child do when he adapts to changes in routine? In what situations do these behaviors occur? What is meant by changes in routine?
	C. Have acquired certain cognitive skills and understandings, including the following:	Cognitive objectives are not stated in behavioral terms.

PROJECT PREPRIMARY Page 3

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy OUTCOMES

Specific Dimensions	Program Definition	Judgments
I. Objectives--changes expected to occur in program participants as a result of the program (contd.)	1. The ability to make better use of sensory perceptions a. Visual b. Auditory. 2. A greater awareness and understanding of people, places, and things a. A larger and more functional vocabulary b. The ability to order, classify, and describe objects on the basis of size, shape, color, smell and touch 3. The ability to use their memories and imaginations to: a. Remember songs and stories b. Express themselves through role playing and fantasy exploration c. Express themselves with such things as paints, crayons, and other art materials	"The ability to make better use of sensory perceptions" might be stated in behavior terms as "The child will demonstrate improved use of sensory perceptions." The kinds of sensory perceptions children will be expected to learn to use more effectively is not spelled out in enough detail.

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy OUTCOMES

Specific Dimensions	Program Definition	Judgments
II. Other Benefits--benefits expected to accrue to other than program participants as a result of the program	Parents of the participants will benefit from the program in the following ways: A. They will acquire a better understanding of the ways in which children develop. B. They will develop an improved image of the school.	
III. Criteria for Successful Completion of or Removal from the Program	Children are generally released after two years in the program. Exceptions can be made for a variety of reasons: A. An extreme inability to deal with the demands of the classroom environment B. Inability to cope with frustration and changes in routine C. An extreme lack of physical or emotional self-control D. An extreme lack of self-confidence manifested in frequent displays of immaturity or hostility E. A marked inability to develop basic cognitive skills as evidenced by an inability to describe ideas or experience in a logical manner, to generalize on the basis of experience, to learn vicariously	

Page 5

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition	Judgments
I. Participants A. Selection Characteristics--the criteria that are used to determine who shall participate in the program	Participants are selected using the following criteria: 1. An unstable or extremely limited family income 2. A one-parent family 3. A large family 4. A one-child family or isolated child 5. A family having three preschool children 6. A working mother 7. Chronic illness on the part of parents 8. Poor housing or crowded living conditions 9. A family history of learning problems 10. Language disabilities or late development of language skills 11. Any other condition that suggests deprivation and indicates the need for compensatory education	How are these criteria applied? Are some more important indicators of need than others? Criteria 1, 8, 9, 10, and 11 should be spelled out in considerably more detail. How little does a family have to be earning before children become eligible for compensatory education? What kinds of language disabilities and learning problems suggest deprivation?

PROJECT PREPRIMARY

Page 6

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition	Judgments
B. Entering Behaviors--characteristics of participants (other than selection characteristics) which are related to performance in the program	1. Physical a. Are generally undernourished b. Have low resistance to disease c. Tire easily d. Are often irritable and/or apathetic 2. Socio-emotional a. Are burdened by feelings of personal insecurity, and have little self-esteem b. Are distrustful of others and reluctant to share objects and affections c. Are impulsive rather than reasoning d. Are extremely adaptable 3. Cognitive a. Have a very short attention span b. Are extremely curious c. Are often unable to express feelings or describe experiences due to a very limited vocabulary d. Have difficulty organizing, classifying, and generalizing	Entering behaviors are not related closely enough to program objectives. Objectives imply something is lacking in the child. This lack should be described in terms of entering behaviors.

PROJECT PREPRIMARY

Page 7

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition		Judgments
II. Staff--qualifications with respect to specific positions			
Staff Member	Professional Qualifications	Personal Qualifications	
Program Coordinators	1. College or university degree in education 2. Teaching experience	1. Patience 2. Adaptability 3. Understanding ✓4. Imagination 5. Ability to work effectively with different kinds of people	The relationship between staff qualifications and functions and duties is not indicated. How much and what kinds of teaching experience are needed?
Teacher	1. College or university degree in education or its equivalent 2. Teaching experience	1. Patience 2. Adaptability 3. Understanding 4. Imagination	
Assistant Teacher	1. Two years of college 2. Experience in working in poverty areas	Same as teacher	Are certain kinds of work experience more immediately relevant and useful for pre-school personnel than others?

PROJECT PREPRIMARY

Page 8

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition		Judgments
Staff Member	Professional Qualifications	Personal Qualifications	
Aide	1. High school education 2. Residence in a poverty area	Common sense	
Volunteers	Participation in training program	Understanding of and interest in children	The duration and content of the training programs are not spelled out.
III. Support A. Administrative Support--administrative personnel who cooperate in carrying out the Preprimary Program	Principals in the individual schools support the program in the following ways: 1. Provide supplementary advice and assistance 2. Support teachers in their dealings with parents 3. Expedite the procurement of necessary materials and supplies		

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition		Judgments
B. Human Resources-non-administrative and non-staff personnel whose contributions and cooperation are necessary to the operation of the program			
Role	Functions	Qualifications	
1. Community Agent	1. Provides a communications link between the school and the home 2. Helps select children using OEO criteria 3. Follows up absences 4. Arranges for trips 5. Provides a male image to whom children can relate both in and out of school	1. Must reside in the neighborhood 2. Should have training and/or experience in social work	

PROJECT PREPRIMARY　　Page 10

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition		Judgments
Role	Functions	Qualifications	
2. Eurhythmics Instructors			
3. Storyteller			
4. Special consultants	Provide medical, dental, mental health, speech, and social work advice and referral services		
C. Media--the materials, supplies, and equipment required for program activities	1. Permanent equipment a. Child-sized tables and chairs b. Cooking equipment c. Balancing boards d. A climbing apparatus e. Housekeeping equipment 2. Educational equipment and materials a. Building blocks b. Hammers and nails c. Beads d. Puzzles and games e. Pictures f. Trucks g. Play dough h. Water play equipment		The list of materials, equipment, and supplies is not complete. The media are not related to specific activities.

Page 11

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy ANTECEDENTS

Specific Dimensions	Program Definition	Judgments
C. Media--the materials, supplies, and equipment required for program activities (contd.)	i. Sand j. Clay k. Chalk and crayons l. Paint and brushes m. Measuring equipment for scientific inquiry and experiments n. Materials for cutting and pasting, and paper in various textures and colors	

PROJECT PREPRIMARY Page 12

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy PROCESS

Specific Dimensions	Program Definition			Judgments
I. Participant Activities--the day-to-day program activities that ultimately lead to the achievement of objectives	The kinds and combinations of activities presently being used in Pittsburgh's Preprimary Program vary as the ingenuity and inventiveness of the program staff. For the purpose of description, however, most of the activities can be categorized as follows:			The list of activities is not complete, and activities are not related to objectives.
	Category	Purpose	Materials or Examples	
	1. Manipulative or constructive	To develop small muscles and improve muscle and eye-hand coordination	1. Puzzles 2. Small blocks 3. Table games 4. Beads 5. Scissors, small hand tools	It is possible for some activities to serve more than one purpose and to be related to a number of objectives.
	2. Motor or large muscle	To develop large muscles and motor skills	1. Large blocks 2. Climbing and balancing equipment 3. Large balls 4. Wheel toys, bicycles, wagons, trains, and trucks 5. Jumping ropes	

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy PROCESS

Specific Dimensions	Program Definition			Judgments
	Category	Purpose	Materials or Examples	
I. Participant Activities--the day-to-day program activities that ultimately lead to the achievement of objectives (contd.)	3. Imitative or dramatic play	To enable children to express themselves through role playing and fantasy exploration	1. A housekeeping corner 2. Doll tub 3. Play dough	An exhaustive list of program activities is beneficial in spelling out program objectives more clearly.
	4. Sensory	To develop sensory perceptions	1. Paints 2. Clay and play dough 3. Sand 4. Water and soap	
	5. Creative experience	1. To facilitate self-expression 2. To develop sensory and aesthetic awareness 3. To promote imaginative thinking	1. Paints, easels and crayons 2. Cutting and pasting materials 3. Modelling clay	

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy PROCESS

Specific Dimensions	Program Definition			Judgments
	Category	Purpose	Materials or Examples	
I. Participant Activities--the day-to-day program activities that ultimately lead to the achievement of objectives (contd.)	6. Cognitive	1. To develop the child's willingness and ability to listen attentively and speak effectively 2. To help the child begin to order, classify, and describe objects in terms of size, shape, and color	All activities in which the child is encouraged to speak correctly and to use language skills to solve problems	
	7. Other activities	To broaden the child's awareness and understanding of people, places, and things in the world around him	1. Snacks 2. Neighborhood walks and trips	

Page 15

PROJECT PREPRIMARY

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy PROCESS

Specific Dimensions	Program Definition		Judgments
II. Staff Functions and Duties with Respect to Specific Positions			
Staff Member	Function	Duties	
Program Coordinator	Provides the overall coordination, guidance, and support needed to plan and implement the program	a. Helps set up the program and plan curriculum b. Orients and trains new teachers and other staff c. Visits schools and provides advice and assistance as needed d. Reads and comments on teachers' logs e. Handles personnel problems f. Audits petty cash reports g. Guides and evaluates case studies of specific children	No attempt has been made to relate functions and duties to program objectives.

PROJECT PREPRIMARY Page 16

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy PROCESS

Specific Dimensions	Program Definition		Judgments
Staff Member	Function	Duties	
Teacher	Plans and implements a series of successful learning experiences	a. Structures an environment conducive to learning b. Provides reinforcement of the learning experience as needed	Teachers' functions are not spelled out in enough detail.
Assistant Teacher	Helps the teacher plan and implement classroom activities		
Aide	Helps the teacher and the assistant under their supervision	a: Prepares materials for various activities b. Performs general housekeeping tasks c. Helps with record keeping d. Prepares snacks and meals e. Assists children under the guidance and supervision of the teacher	
Volunteers	Serves as helper under the supervision of the teacher and her assistants		Duties for volunteers are not specified. Are there any duties which aides or volunteers are explicitly (or implicitly) forbidden to perform?

PROJECT PREPRIMARY

Page 17

REPORT OF PANEL PROCEEDINGS

Section of Taxonomy PROCESS

Specific Dimensions	Program Definition	Judgments
B. Intra-staff Communication and Coordination	1. 'Coordinators' meetings 2. Communications between coordinators and teachers a. Meetings at least once each month b. Exchange of requests, suggestions, ideas, and opinions by telephone or by means of written memos 3. Teachers' meetings 4. In-service training for teachers, assistants, and aides	The frequency and content of staff meetings are not specified. Are certain kinds of problems handled at formal meetings? Are others generally dealt with using more informal means?
C. Communication Between Program Staff and Others	1. Parents and teachers attend weekly meetings. 2. Teachers in some cases attend elementary staff meetings in the schools to which they are assigned. 3. Teachers maintain informal contacts with other elementary teachers, especially kindergarten teachers.	What are the purposes of parents' meetings? Are they important enough to be included in the objectives?

IV. Summary and Implications

A. Overall Comprehensiveness

In order for the Preprimary Program definition to serve as a useful basis for further evaluation of the program, it must be made more comprehensive. For example, objectives are not sufficiently specific and are not stated in behavioral terms. No attempt has been made to distinguish between terminal, ultimate, and enabling objectives.

Once more specific objectives have been established, the antecedent and process dimensions of the definition can be amplified and their relationship to program objectives spelled out in considerable detail. For example, the entering behaviors do not presently describe those characteristics of children which will be modified as a result of their participation in the program and the kinds of deficiencies the program is designed to overcome. The definition is not as specific as it might be in describing other antecedent dimensions including staff qualifications, administrative support, facilities, media, and time constraints and how each of these affects the achievement of program objectives.

Finally, the process dimensions of the definition, the day-to-day activities of the children and the program staff through which human and material resources are combined to produce specified outcomes, are neither complete nor closely related to outcomes and antecedents. The enabling and terminal objectives to which each activity contributes have not been identified.

B. Overall Validity

The Preprimary Program definition as it stands is difficult to assess in terms of internal consistency. Once it has been rewritten in the manner described above, it should be easier to judge the face validity or the reasonableness of the functional relationships set forth.

C. Compatibility

Although Stage I interviews indicate that there have been instances in which the philosophy of the Preprimary Program has created minor problems for some kindergarten teachers, no significant incompatibilities between the program and others being implemented by the Pittsburgh Board of Public Education have been identified. The larger problem of continuity into kindergarten and first grade will be examined in greater detail during the next phase of evaluation.

Appendix A

TAXONOMY OF PROGRAM DIMENSIONS

GENERAL DESCRIPTION	OUTCOMES	ANTECEDENTS	PROCESS
I. Overall Statement of Objectives and Rationale of the Program II. Description of Scope A. Number of Pupils and/or Number of Classes and Schools Involved B. Grades or Ages of Participants C. General Description of Staff	I. Major Objectives A. Terminal Objectives[1] B. Ultimate Objectives[2] II. Enabling Objectives[3] III. Other Benefits[4] IV. Criteria for Successful Completion of or Removal from the Program	I. Students A. Selection Criteria[5] B. Entering Behaviors[6] II. Staff A. Program Personnel by Specific Positions B. Qualifications for Specific Positions 1. Professional 2. Personal III. Support A. Administrative Support B. Human Resources C. Media D. Facilities IV. Time Constraints	I. Participant Activities A. Enumeration of Activities B. Estimate of Time Spent on Each C. Media Used to Carry out Activities D. Activities Related to Objectives II. Staff Activities A. Functions and Duties for Specific Positions B. Intra-staff Communication and Coordination C. Communication between Program Staff and Others and its purpose

[1] Behaviors exhibited by participants at the end of the program which demonstrate successful completion of the program

[2] Long range goals of the program, objectives to which the program hopefully contributes, but for which it does not have sole responsibility

[3] The skills, attitudes, and information which students must acquire during the program to insure the accomplishment of the major objectives

[4] Benefits expected to accrue to other than program participants as a result of the program

[5] Criteria that are used to determine who shall participate in the program

[6] Characteristics of participants (other than selection characteristics) which are related to performance in the program

Appendix B

GROUP INTERVIEW SCHEDULE
PREPRIMARY PROGRAM

1. What are the major objectives of the Preprimary Program? (How does a child benefit from participating in the program?)

2. What are the characteristics of children enrolled in the program? (background, attitudes, knowledge, skills, etc.)

3. What kinds of activities are specified for children in the program? How do these activities contribute to the objectives of the program? Are some more important than others? Why?

4. What criteria do you use to assess the kinds of growth that take place in children as a result of their participation in the program? (changes in attitudes, development of skills, etc.)

5. What are the criteria used to judge a child's readiness for release from the program?

6. What kinds of personnel are needed to carry out the program? (In the classroom? At the supervisory and administrative levels?) How does each contribute to program objectives? What specific tasks does each perform? What qualifications are needed?

7. How do you keep each other informed about purposes, methods, needs, problems, etc.?

Appendix C

PREPRIMARY INTERVIEW SCHEDULE (STAGE I) 1966- 1967

School ______________________________ Date ______________

Respondent ____________________ Interviewer ______________

TEACHER INTERVIEW

1. Are the facilities, equipment, materials, and supplies you have at your disposal adequate to meet the needs of the Preprimary Program? Do you share any facilities, equipment, or materials with other programs?

2. Are qualified personnel generally available when they are needed (substitutes, eurhythmics instructors, medical specialists, mental health specialists, etc.)?

3. Do you and your staff have enough time to plan, prepare, and coordinate classroom activities? If not, how might the necessary time be made available?

4. How sympathetic are school administrators and other teachers to the needs and special problems of the Preprimary Program at ______________________ School?

5. How does the Preprimary Program affect the kindergarten operation in this school?

KINDERGARTEN TEACHER INTERVIEW

1. How does the Preprimary Program affect your Kindergarten Program?

PRINCIPAL INTERVIEW

1. Very generally, how well would you say that the Preprimary Program fits into the overall program of instruction at ____________ School? Have there been any special problems involving the use of facilities or the availability of qualified personnel?

2. How sympathetic are the other teachers to the needs and special problems of the Preprimary Program?

4

Appendix D

RESPONSES OF INTERVIEWEES

Teacher Interview

Question 1

Are the facilities, equipment, materials, and supplies you have at your disposal adequate to meet the needs of the Preprimary Program? Do you share any facilities, equipment, or materials with other programs?

Answers by Respondents

1. Paints and related materials are generally in short supply. There is a need for a slide or something else for the children to climb on (for large muscle development). Share cement play space outside with kindergarten.

2. Art supplies are presently being purchased with funds from the petty cash allotment. Climbing equipment could be put to good use. Share outdoor playground and small basement gym with other elementary classes. Access to gym limited to once or twice each week during most of the year.

3. Facilities, equipment, materials, and supplies are adequate.

4. No equipment for large muscle development. Need more storage space. Share puzzles and games with kindergarten. Outdoor play space not suited to needs of preprimary children.

5. Paper and paints are always in short supply. Share outdoor playground with kindergarten, first grade, and older children.

6. Equipment, materials, and supplies readily available or purchasable with petty cash. Outdoor play space presently under construction. Share demountable with a special third grade which adds to enrichment program.

7. Facilities, equipment, materials, and supplies have been adequate, although more paints and manipulative games would be very much appreciated. The preschool children are free to use the playground at any time and have regular access to the gym.

Question 2

Are qualified personnel generally available when they are needed (substitutes, eurhythmics instructors, medical specialists, mental health specialists, etc.)?

Answers by Respondents

1. Availability of substitutes poses no problem. Would like to have children examined periodically by mental health, medical, and eye and ear specialists. *

2. Eurhythmics program discontinued. A storyteller visits the class periodically.

3. Teacher feels that the Community Agent has too many duties to be able to carry out job responsibilities to the Preprimary Program as well as he should.

4. Qualified personnel have been available when needed.

5. A psychiatrist working with emotionally disturbed children at the elementary school to which the preschool unit is attached has been available for consultation. Unfortunately no provisions for follow up have been made. The eurhythmics program has been discontinued apparently due to a lack of space. Teacher found program helpful but feels she will be able to do as well without it.

6. No ready access to a psychiatrist who could come to observe children who may have emotional problems.

7. Eurhythmics instruction discontinued when classroom formerly used for this purpose was assigned to another program. A psychiatrist has tested a few children who may need another year of preschool.

---

* Shortly after this interview was conducted it was announced that health services would be made available to children in the Preprimary Program by the Allegheny County Health Department.

Question 3

Do you and your staff have enough time to plan, prepare, and coordinate classroom activities? If not, how might the necessary time be made available?

Answers by Respondents

1. Staff has very little time to plan or make preparations. With mother's meetings and other meetings of one kind or another on Fridays, a good deal of the planning must be done on the run or after hours.

2. Occasionally the staff can get together on Friday afternoon for the purpose of planning and coordination. More time might be available if mother's meetings could be held on a biweekly rather than on a weekly basis.

3. Planning is done at lunch and such times as are convenient before and after hours.

4. More time is needed. Planning is sometimes made doubly difficult by the fact that the children may not want to do what has been planned and alternative activities have to be devised on the spur of the moment.

5. There is simply not enough time available for planning. Not sure how extra time might best be made available.

6. Staff members feel that they would have more time to plan if mother's meetings could be held less frequently. As things stand, most of the planning is done by the teacher at home on weekends with little or no consultation with her assistant and aide.

7. In spite of the recent decision to eliminate daily plans and to require only weekly plans, there is still too much paper work to be done. Staff lacks time to better determine needs and interests of individual children.

Question 4

How sympathetic are school administrators and other teachers to the needs and special problems of the Preprimary Program at _____ School?

Answers by Respondents

1. The principal has given the program whole-hearted support since its inception. The other teachers have also been most understanding and helpful.

2. The principal and the kindergarten teacher have been very helpful. Most of the other teachers have been sympathetic although some seem to harbor hard feelings about the teacher-pupil ratio and the materials and equipment that have been placed at the disposal of the preprimary staff. Such feelings are understandable in a situation in which other teachers have large heterogeneous classes and are forced to work with a minimum of materials and supplies.

3. The principal has been very cooperative and has helped me become an integral part of the faculty of the school.

4. The principal has been helpful. Most of the other teachers seem to think three- and four-year olds ought to be at home with their parents.

5. The administrative staff and the other teachers are becoming more sympathetic. At first the latter resented the size of the preschool staff and the facilities and supplies at their disposal. The preprimary teacher also had difficulty keeping the noise level down. This factor added to initial misunderstanding and hostility.

6. Principal helps out when the preprimary teacher needs him. Some of the other teachers may feel that preschool is a baby-sitting operation. Most are interested in what is going on in the preschool.

7. The principal has been very accepting and helps with difficulties such as scheduling. The other teachers complain about the kids from time to time but for the most part are understanding.

Question 5

How does the Preprimary Program affect the kindergarten operation in this school?

Answers by Respondents

1. Hopes that some day something will be done to provide for better continuity between preschool and kindergarten. After two years in the Preprimary Program the children are ready for a more challenging program than the one to which they are presently being exposed in kindergarten.

2. Don't really know except that the kindergarten teacher says the children who have had preschool experience seem to make a quicker adjustment to the kindergarten environment than those who haven't.

3. Because of the sheer size of kindergarten classes and the presence of many children who have not had preschool experience, many former preschoolers are probably marking time in kindergarten.

4. Doesn't seem to have created any special problems so far.

5. There is very little continuity into kindergarten. Until recently neither the preprimary teacher nor the kindergarten teacher knew very much about what the other was doing. Presently visit each other's classes when time permits.

6. Has heard that former preschoolers have adapted fairly well to the routine in kindergarten and that some seem to have done pretty well.

7. Don't know. Kindergarten teacher doesn't seem to understand the purpose and the methods being used in the Preprimary Program.

Kindergarten Teacher Interview

Question 1

How does the Preprimary Program affect your Kindergarten Program?

Answers by Respondents

1. Try to group children according to age and previous experience. Children who have had preschool experience are generally accustomed to greater freedom and more individual attention than they can be given in kindergarten. Nevertheless, they soon learn to adjust to the more formal kindergarten routine.

2. Finds that although children who have attended preprimary are accustomed to considerable freedom to do as they please, they adapt to the kindergarten situation in a relatively short time. Is very much in favor of the Preprimary Program.

3. Former preschoolers are definitely better prepared than children who come straight from home. Difficult to be very specific except to say that the Preprimary Program is a good one and ought to be continued. It would be more effective if the same kind of attention could be provided in kindergarten and subsequent primary grades.

4. The fact that former preschoolers are aware of what is expected of them in school and can serve as an example for those children who have never been in school before makes the kindergarten teacher's job easier. Some parents have complained that their children got too much attention, that they had grown accustomed to having someone help them into and out of their coats and boots and seemed unwilling to begin doing these things for themselves.

5. Most preschoolers have an adjustment of sorts to make when they get to kindergarten. They come used to doing pretty much as they please and have to learn to take turns and to do lots of little things for themselves. As far as learning is concerned, they demonstrate no discernible advantage from having attended preprimary. Don't score higher on Detroit IQ test. To date there has been no transfer of records, so I don't even know who attended and who didn't except from hearsay.

6. Former preschoolers seem to know a lot more but often have trouble getting used to a routine in which they get less freedom

and less individual attention than they are accustomed to. Seem to be a lot more verbal and better adjusted socially than children who didn't attend preprimary. Others don't seem to have learned as much.

7. Feels three- and four-year olds should be at home. Preprimary Program has had the effect of calling people's attention to the importance of kindergarten.

Principal Interview

Question 1

Very generally, how well would you say that the Preprimary Program fits into the overall program of instruction at______School? Have there been any special problems involving the use of facilities or the availability of qualified personnel?

Answers by Respondents

1. Very well. Preprimary teacher needs more storage space.

2. In spite of its physical separation, the Preprimary Program is structurally very much a part of the overall program of instruction.

3. As far as I can tell, the Preprimary Program isn't creating any special problems. The children seem to be getting the attention they need and are enjoying themselves very much. Some need a father image--a male to whom they can relate. Isn't being provided at present.

4. Although the purposes and methods of the Preprimary Program seem sound, it is difficult to see just how it fits into the overall program of instruction at_____School. Perhaps its newness has something to do with the fact that the program stands apart from the others. The teacher seems to have isolated herself from her colleagues even to the extent of sitting by herself at meetings.

5. Isn't sure that the Preprimary Program is accomplishing all of the objectives it set for itself. Economic criteria have been difficult to apply. Feeding program may be unnecessary, at

least in this neighborhood. Children may even be getting the wrong impression of what is expected of them in school. For example, they seem to resent being asked to clean up after themselves when they get to kindergarten. Nor can the preprimary staff hope to provide the love some of the children aren't getting at home.

Question 2

How sympathetic are the other teachers to the needs and special problems of the Preprimary Program?

Answers by Respondents

1. The kindergarten and preprimary teachers share the same building and many materials and supplies. They get along excellently together. The other teachers are very accepting of both the Preschool and the Kindergarten Programs.

2. Teacher opinions are presently being sampled. The kindergarten and other teachers have visited the primary unit on a number of occasions and have come back very impressed. There don't seem to be any unfavorable feelings about the program.

3. There are no real resentments among the other teachers at _______School.

4. Most of the other teachers are passive in their acceptance of the Preprimary Program. Some resent the abundance of materials and supplies at the disposal of the preschool staff.

5. Teachers with large classes resent the teacher-pupil ratio in the Preprimary Program. The recent busing of preschoolers (while other elementary children walked to school) seems to have exacerbated hard feelings and led to misunderstandings between the school and parents.

APPENDIX B

Preprimary Program Problems Inventory

Communication--the biggest problem and most serious threat to the effectiveness of the program.

1. There has been a lack of effective communication about the status, purpose, and methods of the preprimary program among teachers and between teachers and administrators on all levels. As a result, uncertainties and misconceptions about the program continue to burden the preprimary staff and make it difficult to insure continuity into kindergarten and first grade.

2. Understanding and cooperation between parents and staff have been less than satisfactory. Parental interest and involvement in the program has been disappointing.

Administrative Problems

3. Teachers and staff do not feel they have enough time to plan, prepare, and coordinate classroom activities. Too much time is taken up by "paper work" and"burdensome clerical duties."

4. Irregular receipt of petty cash coupled with inflexible restrictions regarding its use make it difficult for teachers to obtain the additional materials and supplies they often need. The suggestion was made that teachers be allotted $50 each month to spend as they see fit.

5. Ordinary materials and supplies are not always available when they are needed, or in the quantities required.

6. In a number of cases community agents have not been well qualified for the tasks they are expected to perform. Perhaps the qualifications of the agent need to be spelled out in greater detail and applied more rigorously in the selection of candidates.

Facilities

7. There is a general lack of outdoor playground space and a shortage of indoor storage space. Provisions for the maintenance (upkeep and repair) of equipment have been inadequate.

Instruction

8. There is some concern about pressures toward greater emphasis on structured learning, especially structured learning of a cognitive nature. This was reflected in differences of opinion regarding the usefulness of detailed planning and record-keeping for each individual child. All participants are willing to be held accountable for what they are doing; most see planning, keeping logs, and standard tests as useful aids which help them become more effective teachers. More specifically, logs were defended as means of making the teacher's objectives and methods explicit, thereby helping to insure consistency and continuity in the preprimary program. They

also serve as a place to record the kinds of progress each child is making as well as the areas in which he may need extra attention.

9. There has been a shortage of qualified substitutes for the preprimary program.

Other

10. Some teachers complained of being unable to secure professional advice and assistance of a medical, dental, psychological, and social work nature when they are needed.

APPENDIX C

Preprimary Program Definition

1. Introduction

Broadly and very loosely defined, the Preprimary Program in the Pittsburgh Public Schools seeks to begin to develop the fundamental socio-emotional characteristics and cognitive skills children need to succeed in school. More specifically, the program is designed to provide three- and four-year olds from economically disadvantaged families with the kinds of emotional security and learning experiences that will enable them to start school on an equal basis with their less deprived peers.

2. Characteristics of Children Enrolled in the Program

a. Physical

Many of the children enrolled in the program are placed at a considerable disadvantage by poor health. They are generally undernourished and often come to school on an empty stomach. As a result, their resistance to disease is low, they tire easily, and they tend to be extremely irritable and/or apathetic.

In spite of these factors, however, most of the children exhibit a measure of physical strength and agility not ordinarily found among three- and four-year olds.

b. Socio-emotional

In making the transition from home to school and in learning to adapt to life in the classroom, children from poor families are subject to a number of socio-emotional constraints. To the extent that they are accustomed to a different kind of parent-child relationship and home environment than their less disadvantaged peers, they bring different attitudes, values, and ways of behaving with them to school.

Most of the children enrolled in the program are burdened by feelings of personal insecurity and often mistrust other children as well as adults. They have difficulty sharing objects and affections. Many are extremely timid and overly cautious. Others are given to displays of extreme hostility such as biting, hitting, or spitting.

Most of the children have very little physical or emotional self-control and tend to be impulsive rather than reasoning. They have difficulty in learning to cope with the kinds of attention and encouragement, not to mention the controls and discipline, to which they are exposed in the classroom.

The willingness of preprimary children to take part in activities which are new to or difficult for them also seems to be conditioned by their image of themselves and the school.

Many three- and four-year olds from poor neighborhoods have little self-esteem and have been exposed to influences that make them anxious about school and even reluctant rather than eager to attend.

Nevertheless, most poor children are extremely adaptable. Once they become accustomed to the classroom environment, they are often better able to get along with each other than many of their more privileged peers.

c. Cognitive

Having had little active encouragement in developing their ability to look and listen or to order, classify, and generalize about what they see and hear, most preprimary children lack the cognitive skills and understandings their more fortunate peers have already begun to develop. As a result, they have difficulty identifying and describing people, things, and places in the world around them. They have very limited vocabularies and express themselves by using fragmented words instead of phrases or sentences. Many have a very short span of attention and, having mostly been talked to rather than with, do not know how to listen attentively. Intellectually they range from dull to exceptionally intelligent. Many are able to make up for their lack of acquired knowledge by means of a very active curiosity.

3. Nature and Purposes of Preprimary Activities

The Preprimary Program attempts to create an environment in which each child can become more aware of himself and many of the people, places, and things in the world around him by taking an active part in a sequence of successful learning experiences. The classroom itself is, therefore, arranged in such a way as to meet the needs and stimulate the interests of three- and four-year-old children. As much as possible, the curriculum is tailored so that each child can do the things he likes best and still be confronted with situations in which he has to find solutions to problems and in so doing, begins to develop new attitudes, values, and habits of mind as well as the cognitive skills needed to order, classify, and describe his experiences.

The kinds and combinations of activities presently being used in Pittsburgh's Preprimary Program are limited only by the ingenuity and inventiveness of the program staff. For the purposes of description, however, most of the activities fall into one or another of the following categories:

a. Manipulative activities involving puzzles, games, pictures, blocks, beads, scissors, and even small tools (hammer and nails) in which the child develops small muscles and eye-hand coordination as well as an awareness of similarities and differences in sizes, shapes, and colors.

b. Motor activities such as climbing, pedaling, balancing, throwing, lifting, pushing with arms and legs, pulling, pounding, and stretching to develop motor skills and large muscles.

c. Dramatic play or imitative activities in which children have opportunities to express themselves through role playing and fantasy exploration. A housekeeping corner including such things as a make-believe kitchen, doll tub, and play dough lends itself to this kind of activity.

d. Activities which enable the child to develop his sensory perceptions and help him to begin to order, classify, and describe objects in the world around him on the basis of size, shape, color, smell, and touch.

e. Creative experiences involving exposure to music, rhythm instruments, drawing, coloring, and painting. These activities enable the child to find satisfying ways of expressing himself, to further develop sensory and aesthetic awareness, to explore and experiment, and to think imaginatively.

f. Activities designed to develop the child's ability to listen attentively and to speak effectively, such as dramatization, verbal repetition, questions and answers, "show and tell," sentence completion exercises, and songs and stories.

g. Other activities including snacks, during which time children learn rudimentary social skills, table manners, and how to use eating utensils; neighborhood walks and trips which broaden children's awareness of people, places, and things around them; and a science table at which children can exercise their natural curiosity and become familiar with the behavior of substances and living things under a variety of conditions.

4. Partial List of Material, Equipment, and Supplies

a. Permanent Equipment. Child sized tables and chairs, storage space, display space, toilet facilities, something to cook on, picture files.

b. Educational Equipment and Materials. Building blocks, hammer and nails, beads, puzzles, games, pictures, balancing boards, climbing apparatus, trucks, dolls and doll house, housekeeping equipment, play dough, cooking utensils, water play equipment, sand, clay, chalk, crayons, paint and brushes, measuring materials for scientific inquiry and experimentation, materials for cutting and pasting, materials with a variety of textures, other objects to feel, paper in various textures and colors.

5. Criteria used to assess the kinds of growth that take place in children as a result of their participation in the program.

Different teachers use different methods to keep track of the kinds of growth that take place in children. Some teachers make a practice of jotting down notes each day; others jot down a word or incident concerning a child; still others make mental notes of their observations. The day's experiences as recorded or remembered are then used as a basis for more detailed observations about each child recorded in the teacher's log. These observations in turn are used to chart the child's progress and to make plans for his further development.

Most of the judgments made by the teachers and aides are based on first hand observations of the children as they participate in various classroom activities. In the area of socio-emotional growth, changes in a child's attitudes toward himself, his peers, and the school environment as reflected in his ability to share objects and affections, the progress he makes from solitary to parallel to cooperative play, and the ways in which he demonstrates self-confidence and initiative are significant indicators of progress.

Growth in language skills can be observed as children begin to use narrative forms, complete phrases, and even sentences to describe sensory experiences. The ways in which young children use new words, precepts, and concepts is a direct measure of the extent to which they understand them.

Recently a series of more objective instruments for diagnosis and evaluation was developed by the Office of Research with the aid of preprimary teachers. One instrument provides a number of yardsticks that can be used to standardize the procedures teachers use to assess the kinds of socio-emotional development taking place in children. The other provides a fairly reliable measure of the extent to which children are learning basic cognitive skills. The usefulness of these instruments should not be exaggerated, however. While they may be fairly reliable for large groups of children and therefore invaluable as a means of assessing the effectiveness of the program as a whole, they are by no means infallible and may be highly unreliable for individuals. Under these circumstances, there can be no substitute for more or less subjective judgments teachers make about the children they work with every day.

6. Criteria used to determine readiness for release from the program.

Children are generally released after two years in the program although exceptions can be made for a variety of reasons: an extreme inability to deal with the demands of the classroom environment or an inability to cope with such things as frustrations and changes in routine. Similarly, an extreme lack of physical or emotional self-control, or a lack of self-confidence manifested in frequent displays of immaturity or hostility may justify keeping a child in the program or transferring him to a special program better tailored to meet his individual needs. Failure to begin to develop the basic cognitive skills may also mean that he will benefit from another year in the Preprimary Program. Teachers mention such things as marked inability to describe ideas or experiences in a logical manner, to generalize on the basis of experience, and to learn vicariously as possible indicators that a child needs special attention.

It should be noted, however, that what is necessarily a relatively subjective judgment is not made by the teacher alone. The advice and consent of parents, other teachers, and professional specialists is always taken into consideration. Nor have any hard and fast rules been established. Each case is decided on the basis of its own peculiar conditions and circumstances.

7. Personnel: Functions, Tasks, and Necessary Qualifications

a. Classroom Personnel

In the final analysis, the success of the Preprimary Program depends upon the effectiveness of teachers and other classroom personnel. They are the axis around which a child-centered program revolves. To understand why three- and four-year olds from poverty neighborhoods think and act the way they do, teachers and their co-workers must understand the environment from which they come, as well as the dynamics of developmental behavior in general. From knowledge and understanding come the qualities of interest and patience that are so crucial if a bond of mutual trust and acceptance between the children and their mentors is to be established and maintained. The staff also emphasizes flexibility, adaptability, and imagination as prerequisites for adults who seek positions in the classrooms. For the purpose of description, the various job categories can be summarized as follows:

(1) The most important single adult in the classroom is, of course, the teacher who is charged with overall responsibility for planning and implementing

a series of successful learning experiences designed to meet the needs and interests of preprimary aged children. The teacher sets the stage or structures the environment in which learning takes place and provides such active reinforcement as may be necessary to expedite the process. In addition to being patient, adaptable, and imaginative, a preprimary teacher must be certified (have a college or university degree in education or its equivalent) and preferably should have some elementary level teaching experience.

(2) The assistant, as the title implies, helps the teacher plan and implement the Preprimary Program in the classroom. Two years of college or university training and some experience in poverty neighborhoods are important qualifications.

(3) The aide is seen as having a function which is ancillary to that of the teacher and her assistant. Duties include preparing materials for the various activities, performing general housekeeping chores, and helping children under the

guidance and supervision of the teacher. Qualifications include a high school education, residence or experience in a poverty area, and a great deal of common sense.

(4) Adult volunteers may be recruited to serve as helpers and housekeepers under the guidance of the teacher and her assistant. Volunteers are expected to understand and be interested in children and to have participated in a brief training program, the purpose of which is to acquaint them with the basic philosophy and methods of preprimary education.

b. Non-instructional Staff

(1) The elementary school principal (when the Preprimary Program is housed in an elementary school) is expected to provide such supervisory and administrative support as may be needed by the classroom teacher and her assistants. Lending moral support, plus help in getting necessary materials and supplies, are seen by the classroom staff as essential parts of the principal's job.

(2) Supervisors provide the overall coordination, guidance, and support needed to plan and implement

the Preprimary Program. They also serve as a communication link between the central administrative staff and the instructors in the schools. The specific tasks performed by the supervisors include helping to set up the program in schools which qualify for compensatory assistance; advising new teachers; visiting schools on a rotating basis (or upon request) for the purpose of offering guidance, expediting purchase and delivery of necessary supplies, and handling personnel assignment problems. As former teachers they are expected to understand the kinds of problems involved in preprimary education as well as the need for constructive suggestions rather than careless criticism of the work being done in the classrooms.

(3) Community agents constitute an important communications link between the schools and the community. More specifically the community agent selects children to participate in the program (using OEO criteria), visits parents to check on absences, explains the purposes of mothers' meetings, and arranges for trips. In many cases

the agent provides a male image to whom children can relate both in and out of the classroom. Ideally, he should have some training in social work.

(4) Specialists in such areas as art, music, eurhythmics, and storytelling.

(5) Medical, nutritional, dental, mental health, speech, and social work consultants.

8. Expediting cooperation and communication.

The people involved in the Preprimary Program keep each other informed about the purposes, methods, needs, and problems related to the work they are doing by getting together formally and informally at frequent intervals.

Supervisors meet with teachers at least once each month. From time to time they also exchange requests, suggestions, ideas, and opinions over the telephone or by means of written communications.

Teachers and parents meet weekly to discuss the program and problems the children are having, as well as the kinds of progress they are making.

Teachers and in some cases their assistants attend preprimary teachers' meetings and, occasionally, elementary staff meetings at the schools in which they work.

APPENDIX D

RESPONSE FORM
PREPRIMARY PROGRAM

1. Your relationship to the program (e.g., teacher, aide, school principal, etc.):

2. Your reaction to the definition meeting held on either February 10 or February 24: (If you did not attend either meeting, please indicate question not applicable.)

3. Your reaction to the program definition issued by the Office of Research:

4. Your reaction to the Preprimary Program (e.g., Do you believe it will work? Is it a worthy effort?):

APPENDIX E

Preprimary Program Observations Guidelines *

Name of Child ____________________

Age (at time of observation) _______

Date and Time of Observation______

Observer's Name__________________

I. Description of Setting

A. Physical environment (draw shape and location of facilities if possible)

B. Preliminary arrangements by the teacher (note any daily modifications)

C. Other information about the situation in which the observation begins

II. Observations (focus on the play activities of individual children)

A. What kinds of play did the child become involved in? How did the activity begin?

B. What materials and equipment were used and for what purpose? What kinds of materials and equipment did the child have to choose from?

C. What kinds of interruptions took place? How did the child respond?

D. How intensively did the child become involved in what he was doing? (Note physical, facial, verbal, and other manifestations of involvement.)

E. What kinds of interactions took place between the child and other children?

* Adapted from Nancy Trevorrow Carbonana, Techniques for Observing Normal Child Behavior (Pittsburgh: University of Pittsburgh Press, 1961).

F. What kinds of interaction took place between the child and the teacher and/or other adults? What was the teacher's or other adult's role in relation to the activity? (Describe physical, facial, and verbal manifestations of direction, reinforcement, or control, both positive and negative.)

G. How long did the activity last? How was it terminated?

III. Interpretation (incorporate follow-up interview with teacher as needed)

A. What did the child seem to be getting out of the activity?

B. What role did the adults play?

C. How were the child's individual needs and interests recognized?

D. How were his responses, ideas, and other contributions utilized?

E. Note anything else which you feel was of special significance.

APPENDIX F

Socio-Emotional Scales and Directions for Use

Activity Area

1. Large muscle - jungle gym, bicycles, etc.
2. Creative - art activities
3. Dramatic play, etc.
4. Manipulative or constructive - table games, pegs, puzzles, blocks
5. Sensory - clay, sand, water
6. Cognitive - vocabulary, discrimination, rhyming, etc.

Coded Scales

I. Initiates Own Activities
   a. No activity or destructive activity
   b. Participates in activity at direction of the teacher/or other children
   c. Requests specific activity
   d. Persues activity on his own initiative

II. Relates to Group and Participates in Group Activities
   a. Non-communicative with others in Group
   b. Silent, solitary participation
   c. Relates to child/teacher - (may be aggressive behavior)
   d. Relates to and becomes integrated into the group

III. Stays With a Task Until Completion
   a. Over-under stimulated - short attention span - random activity with no particular goal
   b. Briefly attempts activity
   c. Channels activities toward a particular goal
   d. Stays with a task until completion

IV. Awareness of Changes in Environment
   a. No apparent notice of new objects of materials in the room
   b. Shows only slight interest in and awareness of new objects or materials
   c. Spends time manipulating, examining or watching unfamiliar objects or material
   d. Asks questions about or uses unfamiliar objects and materials

V. Shows Trust by Accepting School Environment
   a. Enters room only by force
   b. Enters reluctantly - needs encouragement
   c. Enters room - stands on sidelines
   d. Enters room willingly

VI. Shows Trust by Sharing Objects and Affections
  a. Rejects objects and/or attentions of others
  b. Refuses to relinquish objectives and/or attention of others
  c. Shares objects and/or attention of others with limitation
  d. Shares objects and/or attention of others freely

VII. Controls Emotions in Frustrating Situations
  a. Physically aggressive reaction to frustration - hitting, biting, excessive crying, etc. - or complete withdrawal
  b. Responds with verbal rather than physical aggression or withdrawal to frustration
  c. Begins to accept restraints without emotional display
  d. Controls emotions, accepts restraints and criticism

VIII. Adapts to Changes in Routines
  a. Emotional display over rescheduling of activities
  b. Shows general disorientation
  c. Upset by change, but accepts it
  d. Adapts to changes in scheduled activities with ease

IX. Takes Care of Personal Needs
  a. Does not dress himself, toilet, feed himself, etc.
  b. Handles dressing and toileting, feeding with some help
  c. Dresses/toilets/feeds himself
  d. Takes care of all personal needs (anything over and above dressing, feeding toileting)

Directions for using social scales:

1. In order to crystalize your thinking, the activity areas have been divided into six categories.

2. The nine areas of behavior are divided into four steps.

3. The chart which follows is numbered according to the area of behavior and the four steps involved.

4. For each of the first four areas of behavior, put the number of the activity area on the step where it belongs.

	a	b	c	d
I.	1, 3	2	4, 5	

Suppose for "initiates his own activities" he falls on step a for large muscle and manipulative activities. Enter the numbers 1 and 3 in step a.

Suppose he falls on step b for creative area. Enter 2 on step b.

Suppose he falls on step c in the cognitive and group activities. Enter 4 and 5 on step c.

5. For the behavior areas V-IX it is not possible to designate by activity area. Enter the behavior by blackening the space of the proper step.

6. If you find that the progression between steps is too large, please define the steps you feel should come in between.

7. Make out one chart for each child in your class.

Below are listed the activities and materials that would be included under each activity area. These should serve to further clarify the activity areas.

1. Large Muscle -	Hollow blocks or large unit blocks Any climbing equipment Balancing equipment Large balls Bodily coordination - running, skipping, rolling,etc. Wheel toys - bicycles, wagons, trains, trucks, wheelbarrow, baby buggies
2. Art Activities -	Easel, crayons, cutting and pasting Clay if used for creating pre-conceived object (This category does not include finger paint)
3. Dramatic Play -	Role playing in any area, such as housekeeping, blocks, props., etc.
4. Manipulative or Constructive -	Puzzles Small unit block Table games Beads Scissors - used solely for manipulative purposes
5. Sensory -	Finger paint Play dough Clay Sand Water Soap
6. Cognitive -	Use of language Investigative activity - books, science table Discrimination - Also see cognitive objectives

APPENDIX G

The Preschool Cognitive Test

Outline Description

I Vocabulary

I-A Vocabulary--Nouns
Objective: Names and identifies nouns
Test Item: Shown a series of 10 pictures of nouns
Direction: What is this? (for each picture)

I-B Vocabulary--Verbs
Objective: Names and identifies verbs
Test Item: Shown a series of 10 pictures depicting action
Direction: What is the boy (or girl) doing?

I-C Vocabulary--Adjectives
Objective: Names and identifies adjectives
Test Item: Shown a series of colors and of pictures
Direction: For colors--What is the name of this color?
For pictures--Show me something that is slow.

II Generalization and Classification

II-A Generalization
Objective: Selects several differently depicted objects as the same object
Test Item: Shown three different chairs and a bed
Direction: Show me all the chairs

II-B Classification (general categories)
Objective: Selects objects belonging to the same category
Test Item: Shown three fruits and a vegetable
Direction: Show me all the fruits

II-C Classification (uses)
Objective: Selects objects used for same purpose
Test Item: Shown three musical instruments and a watch
Direction: Show me all the things that are used to play music

II-D Relationships
Objective: Finds relationships between objects
Test Item: Shown a picture of a saucer, block, kite and cup
Direction: Show me what goes with the saucer

III Visual Discrimination

III-A Visual Discrimination--Concrete objects
Objective: Makes discriminations between concrete objects
Test Item: Shown series of 4 pictures, two of which are identical
Direction: (Pointing to first picture) Show me the picture that is just like this one

III-B Visual Discrimination--Abstract shapes
Objective: Makes descriminations between abstract shapes
Test Item: Shown series of 4 abstract shapes, two of which are identical
Direction: (Pointing to first picture) Show me the picture that is just like this one

III-C Visual Discrimination--Letters
Objective: Makes discrimination between letters
Test Item: Shown series of 4 letters, two of which are identical
Direction: (Pointing to first letter) Show me the letter that is just like this one

III-D Visual Discrimination--Words
Objective: Makes discriminations between words
Test Item: Shown series of 4 words, two of which are identical
Direction: (Pointing to first word) Show me the word that is just like this one

III-E Visual Discrimination--Memory
Objective: Identifies a previously seen object, shape, letter and word
Test Item: Shown a picture (or shape, letter, or word) on one page and then shown series of four pictures on another page, one of which is identical to first picture

Direction: Shown the first picture: Look at this
Then shown next page: Now find one just like it here

IV Auditory Discrimination

IV-A Auditory Discrimination
Objective: Selects picture corresponding to spoken words
Test Item: Shown series of three pictures having similar names (e.g., pin, pan, pen)
Direction: Find the pen

IV-B Auditory Discrimination--Oral Rhymes
Objective: Supplies rhyming word
Test Item: Two line rhymes repeated by examiner (e.g., I'm thinking of something that rhymes with fair; It's big and it growls. It's called a _______.
Direction: Listen. Finish this rhyme

IV-C Auditory Discrimination--Rhymes
Objective: Selects picture that rhymes with stimulus picture
Test Item: Shown series of four pictures, two of which have rhyming names
Direction: All pictures are named; (pointing to first picture) What rhymes with ring

IV-D Auditory Discrimination--Beginning Sounds
Objective: Selects picture depicting objects beginning with same sound
Test Item: Shown sample, and three other pictures, one of which begins with same sound as sample
Direction: All pictures are named; (pointing to first picture) What begins with the same sound as monkey

V Comprehension

V-A Comprehension--2-picture sequence
Objective: Orders a 2-picture series by selecting picture depicting what happened first in story

Test Item: Shown two pictures and given description of each picture
Direction: If we wanted to put these pictures in order so the story would make sense, which picture would come first?

V-B Comprehension--3-picture sequence
Objective: Orders a 3-picture series by selecting picture depicting what happened first in story
Test Item: Show three pictures and given description of each picture
Direction: If we wanted to put these pictures in order so the story would make sense, which picture would come first?

V-C Comprehension--Problem Solving
Objective: Indicates solution to a problem by selecting picture to finish a sentence (e.g., Shown picture of grocery store, library, money, museum)
Direction: Finish this sentence. If we want to buy meat, we go to the _______.

VI Mathematical Readiness

VI-A Patterning
Objective: Selects picture to extend a pattern sequence
Test Item: Shown a pattern of alternating colors, shapes, or both
Direction: Shown three alternatives; Which one of these comes next in the pattern?

VI-B Counting
Objective: Counts 10 shapes
Test Item: 10 shapes
Direction: Count these circles and tell me how many there are all together

VI-C Identification of Shapes
Objective: Can identify a circle, square, triangle and rectangle

Test Item: Shown picture of circle, square, triangle and rectangle
Direction: What is the name of this shape? (for each shape)

VI-D Comparison of Sets
Objective: Selects the set with more objects
Test Item: Shown two sets of objects
Direction: Show me. Which has more?

VII Directional Spatial Relationships

VII-A Directional Relationships
Objective: Indicates understanding of directional-spatial concepts by selecting picture corresponding to verbal description
Test Item: Shown series of three pictures, each one showing an object in a different position (e.g., book on top of chair, book under chair, book beside chair)
Direction: Show me the picture that shows the book is under the chair

VII-B Opposite Analogies
Objective: Supplies word to complete opposite analogy
Test Item: Listens to sentence repeated by examiner (e.g., wood is hard, a pillow is ______)
Direction: Finish this sentence with an opposite

VII-C Recognition of Rotated Objects
Objective: Identifies picture of an object rotated in space to match stimulus picture
Test Item: Shown sample, and four alternatives, one of which is same as sample but has been rotated in space
Direction: (Pointing to sample) See this. Find one here (pointing to alternatives) that is just like it

APPENDIX H

Preprimary Evaluation
Preliminary Plans, Cycle II

Program Dimension(s)	Question(s)	Rationale for Question(s)
All dimensions of the initial specifications, with special attention to those dimensions identified by the Stage I Panel.	Have change decisions about program design been made and implemented? How can these best be incorporated into the written definition? Or, does the initial definition remain accurate as a description of the current consensus?	As the baseline specifications for evaluation purposes, the written definition of the program must reflect the current consensus of program staff.
All dimensions of the revised, or current, program definition.	Is the current definition comprehensive and internally consistent; is the program as currently defined compatible with other programs in the total school environment?	If the definition has been revised or modified, Stage I criteria should be reapplied, if not revised or modified, negative judgments must be reviewed.
Antecedent dimensions, current definition: a. Entering behavior	What are the social and emotional characteristics of children entering the program at each age level? Are socio-emotional ratings of individual children (using the socio-emotional scales) by different raters consistent?	To determine whether the operating program is congruent with the defined program (i.e., whether assumptions about levels of entering behavior are accurate.) To determine the reliability of the socio-emotional scales used for measuring entering behavior.
b. Support mechanisms and procedures	What is the status of administrative support and supplementary resources in the operating program?	To determine whether the operating program is congruent with the defined program (i.e., whether support needed and resources needed are indeed available.)
Process dimensions, current definition: a. Teacher functions and duties	Do teachers interact with children in the manner prescribed by the program definition?	To determine whether the operating program is congruent with the defined program.
b. Intra-staff communications	Are the specified communications channels extant and operating?	To determine whether the operating program is congruent with the defined program.
Outcomes: Acquisition of cognitive skills	How do disadvantaged children compare with their non-deprived peers in the mastery of basic cognitive skills as measured by the Preschool Cognitive Test?	To continue the development of the preschool cognitive test

Preprimary Evaluation (contd.)
Preliminary Plans, Cycle II

Source(s) of Information	Instruments	Proposed Analysis	Provision(s) for Feedback
Program director and/or Entire program staff (or sample)	Unstructured interview by evaluator using check-list relative to change activity, developed by evaluation staff Group interview schedule and working formats for collecting data on program dimensions, by evaluation staff	Content analysis followed by synthesis	Revised definition mailed to program staff with response forms for return of comments and criticisms Informal contacts between evaluator and program staff for gauging staff reaction to the revised definition
Current definition and panel of judges	Panel meeting, using worksheets. Arrangements made and materials prepared in advance by evaluators	Content analysis of panel proceedings Summary of findings	Report of panel findings mailed immediately to program staff and reported again in Cycle II Evaluation Report
Paired ratings on random samples of children for each age level by professional raters both from within and outside the program	Socio-emotional scales previously developed by collaboration between program staff and evaluation staff	Descriptive summary of data by sample. Relate parameters thus depicted to expectations expressed in program definition Correlation of paired ratings to assess reliability of scales	Informal feedback to raters on findings relative to scale reliability Report of modifications and of congruence or discrepancy of antecedent dimensions in Cycle II Evaluation Report
Field staff opinion and description of relevant operations	Interview schedule by evaluator	Content analysis of interview responses and comparison with corresponding specifications in the definition	Informal feedback to raters on findings relative to scale reliability Report of modifications and of congruence or discrepancy of antecedent dimensions in Cycle II Evaluation Report
Operating program as seen in the preschool classrooms	Observation checklist developed by evaluation staff using data from preliminary observations	Descriptive summary and comparison with specifications in the program definition	Report of congruence or discrepancy of process dimensions in Cycle II Evaluation Report
Perception of the program by school personnel outside the program	Interview schedule by evaluator	Content analysis of interview responses. Inferences about the operation of communications channels based on perceptions of non-program staff	Report of congruence or discrepancy of process dimensions in Cycle II Evaluation Report
Samples of preschool program population, kindergarten population and first graders with and without experience in the program	Preschool Cognitive Test being developed by evaluation staff	To be determined	Findings to be included in Cycle III report or annual evaluation report, whichever is earlier

Chapter 3

A COMPENDIUM OF EVALUATION CASE STUDIES

Case Study I

A NARRATIVE OF PROGRAM EVALUATION
TITLE I, ELEMENTARY AND SECONDARY EDUCATION ACT SERVICES
CINCINNATI PUBLIC SCHOOLS, 1966 AND 1967

Elementary and Secondary Education Act program evaluation in the Cincinnati Public Schools is defined as the attempt to assess the collective impact on the pupils of the various services provided by the Education Act. By contrast, project evaluation is aimed at assessing the impact of a discrete set of services defined as a project. This report provides a brief description of the procedures used in program evaluation during the first year and the approaches and procedures planned for the present year. Evaluation of ESEA program and projects is the primary responsibility of the Division of Program Development (a unit of the Department of Instruction), James N. Jacobs, Director, and Joseph Felix, Associate. Associate responsibilities are assumed by the Division of Evaluation Services, Joan Bollenbacher, Director; the Division of Psychological Services, Charles Miller, Director; and the Division of Data Processing, Edward Ebel, Director. For a detailed report of the first year of Education Act Program Evaluation, the reader is referred to the Journal of Instructional Research and Program Development, Volume 2, No. 1, October, 1966, published by the Department of Instruction of the Cincinnati Public Schools.

Description of Total ESEA Program. Title I of the Elementary and Secondary Education Act of 1965 resulted in an allotment of three million dollars for the Cincinnati Public Schools to enhance the education of disadvantaged children. Thirteen projects were designed for the first year (second semester, 1966). These have been recombined into six projects for the present school year, 1966-67, but the services have remained essentially the same. The projects are listed below with brackets indicating the combinations for the present year.

Present Year 1966-67	First Year 1966	PROJECT
1	1	Early Childhood Education
2	2	Health Services
3	3	Communications (Speech Improvement)
	4	Emotional Disturbances and Learning Disabilities
4	5	Elementary Remediation and Enrichment
	6	Saturday Morning Enrichment Classes
	7	Elementary Summer School
	8	Educational Resource Centers (first year only)
---	9	Parent Education (divided between Elementary and Secondary Projects during second year)
5	10	Secondary Remediation and Enrichment
	11	Secondary Summer School
6	12	Staff Development
	13	In-Service Training

Two classifications of schools were identified to receive Education Act services. The first, primary target schools, are those which have the highest concentration of disadvantaged children. (The proportion of pupils from families on welfare was the criterion used in selecting target schools.) All ESEA projects were operated in this group. The second, secondary target schools, are those with fewer disadvantaged pupils and where only certain projects were operated. A third group, "control" schools, were designated for evaluation purposes. These were the schools which were closest to the target schools in terms of concentration of the disadvantaged. The numbers of types of schools in each classification for the first year were as follows:

Primary Target (PT)	Secondary Target (ST)	Control (C)
13 public elementary	19 public elementary	3 public elementary
3 non-public elementary	14 non-public elementary	2 public secondary
4 public secondary	4 public secondary	

Evaluation Problems. The difficulties inherent in program evaluation were recognized at the outset. The projects operated for only five months of the first school year, obviously too brief a time to detect measurable advantages. The difficulty in attempting to attribute cause and effect relationships to the various projects was apparent, since each service is presumed to interact with every other service. Further, the types of experimental designs required for assessing cause and effect were not applicable, since eligible pupils could not be denied the services for the sake of serving as controls. In addition, even if the above difficulties could have been resolved, it was recognized that it is unrealistic to assume that one cause will produce a given effect, since there are many factors which affect complex variables such as achievement.

First-Year Evaluation Strategy. The overall strategy for program evaluation involved the identification of several complex variables which were viewed as barometers of educational "health." These variables are assessed under eight headings: (1) Teacher Evaluation; (2) Student Evaluation; (3) Parent Evaluation; (4) Pupil Achievement; (5) Pupil Self-Image; (6) Pupil Promotion; (7) Pupil Attendance; and (8) Pupil Drop-Outs.

Data to serve as criterion measurements in each of the eight areas were obtained for the primary and secondary target schools, as well as for the control schools. The general hypothesis which permeated the program evaluation is that the criterion measurements would respond to the intensity of treatments. (That is, $H : PT > ST > C$.) Four criterion measurements were baseline data only, while the remaining four were represented either by pre-, post-test data or baseline data established in previous years. Comparison of PT, ST, and C were ambiguous when only baseline data were available. For example, while "control" schools were chosen for their similarity to the target schools, they could not be assumed to provide a comparable population. Obviously, had the criterion measurements not been higher, it is likely that the schools would have been in the target area.

Change scores were examined on four of the criterion measures making comparisons of PT, ST, and C less ambiguous. This strategy, however crude, did indicate some revealing differences which supported the general hypothesis.

Evaluation Instruments. The matter of instrumentation was no small problem. Some instruments had to be designed locally; others were available from commercial publishers. The concerns over appropriate levels of reliability and validity were ever present. Technical considerations were only part of the problem, however. The impact on staff morale and public relations resulting from administering staff, pupil, and parent surveys could not be overlooked. The upset in the schools involved in scheduling and administering numbers of surveys and achievement batteries was considerable.

Achievement Tests from the Stanford Battery were given in grades 3 through 11. This battery was selected because it was used in the city-wide program with some data available from past years for comparison purposes. The Metropolitan Primary Test was given in grade 2 because of its earlier use. (The Stanford is being given during the present year for ESEA evaluation.)

Survey Instruments were constructed to provide data regarding teacher, pupil, and parent opinions and pupil self-image. A brief description of each follows:

Teacher Survey. The semantic differential technique was used in constructing this survey which consisted of 44 items, each representing a concept or service relating to one or more objectives of the thirteen projects. Teachers rated each item from "poor" to "good" on a 7-point scale. The instrument was printed on a Digitek form which was reproduced locally by the Multilith process. The survey was administered in January and June.

Pupil Survey. The questionnaire consisted of 20 items to which pupils gave "yes" or "no" answers. The instrument was designed to elicit information regarding pupils' attitudes and concerns about school, their parents' involvement in the pupils' education, pupils' aspirations, and their participation in specific activities relating to ESEA project objectives. This survey also was reproduced locally in the Digitek format.

Parent Survey. The parent survey was conducted by interviewing a random sample of 20 parents in each target and control school area. The interviewer read a 14-item questionnaire to which each parent responded "much," "some," or "not at all" to each item. The items related to the parent's opinion of the child's attitude and interest in education and to the parent's opinion of the school and its effectiveness. The Digitek format was used.

Pupil Self-Image. Three instruments were used in the attempt to evaluate pupil self-concept: What I Am Like, Attitudes Toward Self and School, and the House-Tree-Person Test. Each was administered to a random sample of pupils.

What I Am Like. This instrument, developed locally, was based on the semantic differential technique in which pupils in grades 4 through 11 rated themselves on a 5-point, bipolar adjective scale. It consisted of three subtests of 10 items each. The first, What I Look Like, consisted of adjectives characterizing physical attributes; the second, What I Am, attempted to measure self-image from a psychological point of view; the third, What I Am Like When I Am with My Friends, concerned social attributes. Again, the Digitek format was used.

Attitudes Toward Self and School. The Attitudes Toward Self and School or "Faces" test consisted of 18 items, each having two faces, one smiling and one frowning. The format was Digitek, and the pupils in grades 1 through 3 were asked to blacken the nose of the picture that described how they felt about a given question. For example, "How do you feel when you get your report card and take it home?" If a pupil marked the smiling face, it was assumed that he had a positive attitude. If he marked the frowning face, it was assumed that his feelings were more negative.

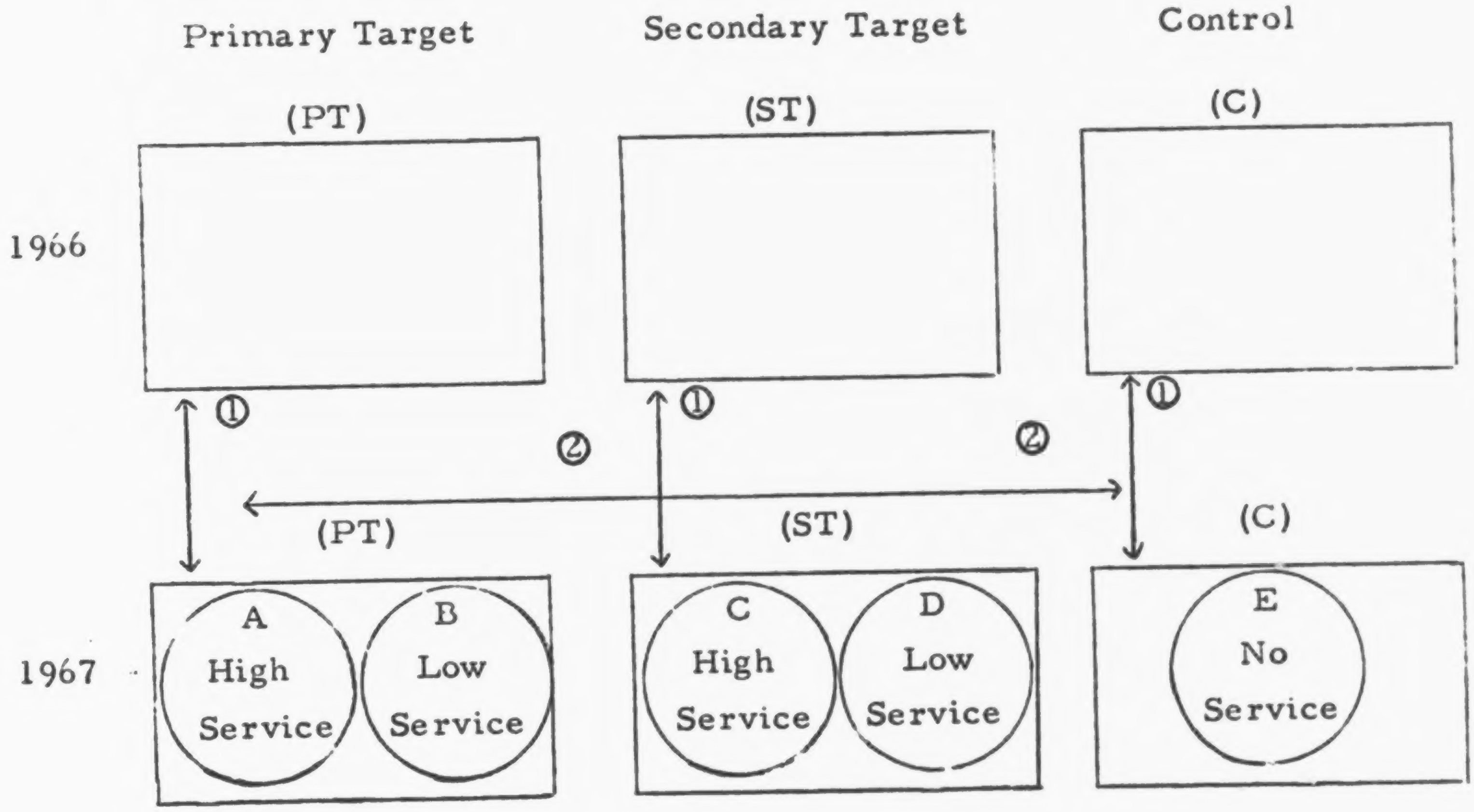
House-Tree-Person. This test involved a projective technique in which pupils drew a picture of a house, a tree, and a person. The tests were scored on eight factors which seemed to have the support of experimental evidence and clinical cross-validation by more than one author. A scoring system was developed to measure the degree of presence of each factor. Each factor was scored on a 3-point scale, making the maximum score 24.

First Year's Report. The first year's ESEA program evaluation was oriented primarily toward establishing baseline data on several important dimensions relating to the educational "health" of the target schools. Details are reported in the October 1966 issue of the Journal of Instructional Research and Program Development mentioned earlier. For this narrative, suffice it to say that the baseline data will serve as a point of departure in the measurement of hoped-for gains in the future ESEA program evaluation in the school system. Further, it is hoped that such baseline data will represent a kind of challenge to staff members--that is, the data will be looked upon as records to be broken.

Program Evaluation Strategies, 1966-67. The criterion measurements used last year will be obtained again this year during April and

May. While the evaluation strategies last year used only the school as the sampling unit, this year's evaluation also will use pupils as sampling units. This strategy evolved from the situation within the target schools where some number of pupils receive intensive service. This situation permits comparison of high service and low service (operationally defined) pupils within the same school. These pupil comparisons, matched on pre-achievement, sex, and grade, should be more sensitive to differences than school-wide averages which tend to conceal possible differences in pupils who receive differential treatment.

The diagram on the following page shows the three types of comparison strategies which will be used during the current year.


Primary Target
Secondary Target
Control
(PT)
(ST)
(C)
1966
①
①
①
②
②
(PT)
(ST)
(C)
1967
A
High
Service
B
Low
Service
C
High
Service
D
Low
Service
E
No
Service

Comparison 1. (Vertical arrows)

Determination of significant change from 1966 to 1967 within PT, ST, C. $H : PT_{67} > PT_{66}$, etc.

Comparison 2. (Horizontal arrows)

Determination of significant change among PT, ST, and C

$H : PT_{increase} > ST_{increase} > C_{increase}$

Comparison 3. (Circles)

Determination of significant differences in change scores among A, B, C, D, and E groups

$H : A > C > B > D > E$

Case Study II

EVALUATION OF THE MOBILE READING CLINIC PROGRAM

DADE COUNTY PUBLIC SCHOOLS, FLORIDA

1965-1966

INTRODUCTION

The public school system of Dade County, Florida, was allotted a sum of $3, 528, 826. 16 from Public Law 89-10 funds to implement programs to serve educationally deprived students with instructional programs or services. Various programs were initiated during the 1965-66 school year and the summer school of 1966. Eight projects were designed and put into effect:

1. Visiting Teacher Counselor Project
2. Special Reading Teachers
3. Mobile Reading Unit
4. Guidance Services Project
5. Junior High Learning Laboratory
6. Special Education Project
7. Classrooms for Headstart Project
8. Summer School Program

As required by the law, the degree of effectiveness of each project was assessed. The description of each project contained a section stating how it would be evaluated. For this purpose, evaluation was interpreted as being concerned with the amount of change induced by a project.

Description of the Evaluation Unit

To accomplish the task of evaluating Federal Projects in Dade County, an evaluation team composed of a professional staff of one coordinator and three special teachers was formed. The titles and specific duties of the evaluation team are as follows:

Coordinator

A. Assume responsibility for carrying out the tasks assigned to the Unit within the general area of evaluating special programs for educationally disadvantaged children

1. Design procedures for
    a. Gathering pre- and post-project data
    b. Maintaining records
    c. Preparing reports

2. Aid in selection and preparation of needed instruments for measuring achievement of stated project objectives

3. Assist in checking the validity and reliability of data-gathering instruments and procedures

4. Determine that appropriate procedures and techniques for the processing and analyzing of data are conducted

B. Supervise the work of the other members of the Unit

C. Evaluate the internal effectiveness of the Unit and make recommendations for improvement in organization and operation of the Unit

D. Work with the managers of special projects, the director of special programs, and other interested school system personnel in planning programs for disadvantaged children and youth

Specialist--Research Methodology and Statistics

1. Work with other members of the Unit in conducting evaluation projects

2. Give special attention to the research design methodology and statistical analysis procedures in the evaluation projects

3. Perform such other duties as may be assigned by the Coordinator of the Evaluation Unit

Specialist--Tests and Measurements

1. Identify suitable instruments by which objective and subjective measurements may be obtained

2. Establish a working rapport with building representatives responsible for testing, Assistant Principals, and Test Chairmen

3. Assume responsibility for seeking out additional instruments and methods, with special attention to the selection, creation, and validation of data gathering devices, and assist in the assessment of the Project objectives

4. Inform the members of the Evaluation Unit of new instruments, recent developments, and current research in the field of tests and measurements

5. Work with other members of the Unit in conducting evaluation projects

6. Perform such other duties as may be assigned by the Coordinator of the Evaluation Unit

Specialist--Research and Information

1. Assist with the implementation of the instruments

2. Process gathered data and produce the desired analytical and computational documentation

3. Assume responsibility for the maintenance of liaison between the Evaluation Unit and the available data processing facilities

4. Contribute to the development of electronic procedures that will automate functions of the Evaluation Unit

5. Work with other members of the Unit in conducting evaluation projects

6. Perform such other duties as may be assigned by the Coordinator of the Evaluation Unit

Initiation of a Project

The Dade County Public School System has established procedures to permit the implementation of special programs into its ongoing operations. An Assistant Superintendent for Special Programs has been appointed by the Dade County School Board to coordinate the functions of special K-12 programs. The Assistant Superintendent has a staff of experts who assist with writing and management aspects of special projects.

It is recognized that ideas for projects and special programs may come from a number of sources, with two such sources being (1) an idea that an individual educator might have or (2) a more universally accepted idea possessed by many who are responsible for educational programs in a large school system, such as preprimary types of schooling. The initiator of an idea for a special program prepares an abstract of what he anticipates if the project should be undertaken in Dade County. A committee composed of five members, with the Assistant Superintendent for Special Programs as the chairman, examines each abstract to determine the program's appropriateness for Dade County.. If a favorable report comes out of the committee's evaluation of the idea in abstract form, the Assistant Superintendent for Special Programs becomes the initiator of the idea. It is his responsibility to see that the project is written within appropriate guidelines established to secure the funding. When a project is in its final form, copies are distributed to the Dade County School Board members for their endorsement. A resolution that they support the project is acted upon in an open School Board meeting.

The Title I of Public Law 89-10 projects of Dade County came into existence after having gone through the procedure just described.

Since Dade County projects under Title I, Public Law 89-10, were accepted and implemented in fiscal 1966, their duration has probably been too brief to expect significant measurable results. Each project had to be superimposed on an already existing instructional program, and achieving efficient operation was thus made a more complex task. Services that had direct impact on individual pupils did so for a very short period of time. Some of the objectives that were identified for projects were so complex in nature that it was not realistic to assume or to expect changes after such a short period of time.

Test Results

When considering standardized test results it must be understood that different forms and norms of the tests were used for pre- and post-testing. Test items on the pre-test were different from those on the post-test. This necessitated the converting of all mean scores to a standard "T" score for comparability. In the evaluation phase of this report, all mean scores are reported as standard scores.

The "T" score mentioned above is based on the percentile rank of the mean score of the group tested with respect to the national norm. A mean change of zero (i.e. the mean score on the pre- and post-tests are the same) does not indicate zero growth, but rather, the educational gap existing between the disadvantaged students and the national norm was the same after the treatment as it was before the treatment.

MOBILE READING CLINIC

Objectives

1. To identify and diagnose students with severe reading disability and to provide remedial instruction for those pupils unable to function in classroom or corrective reading programs
2. To improve student attitudes towards school (reflected by improved attendance and conduct)
3. To improve student performance in both course work and on standardized achievement tests, as a residual effect

General Description

One of the projects initiated this past year was to establish a Mobile Reading Clinic to conduct an extensive and intensive clinical reading program for boys and girls having an urgent need for clinical diagnosis and instruction with their reading problems. Five reserved school buses were remodeled and equipped to serve needy pupils. The bulk of the remodeling was conducted through the Dade County School Board Maintenance Department. Separate compartments with storage facilities were included in the buses. Since five mobile centers were equipped to serve the clinical needs of pupils having severe reading problems, fifteen clinicians were employed to handle the buses, with three clinicians assigned to each bus. The mobile centers were stored in secure compounds during the evening hours. Each center visited two schools a day. Since this program commenced as late as February 1966, only two schools were served by each unit, with the intention of bringing the clinical services to the door of schools having students with severe reading problems. The following ten schools were served by the Mobile Reading Center:

Bunche Park Elementary
North County Line Elementary
Miramar Elementary
Highland Park Elementary
Holmes Elementary
Orchard Villa Elementary
A. L. Lewis Elementary
West Homestead Elementary
Moton Elementary
F. C. Martin Elementary

The Mobile Reading Center Clinic was designed to work with boys and girls ranging from ages 8 through 12, in grades three through six. Three hundred ninety pupils were diagnosed in the short span of the operation of this program. Of this number of pupils, two hundred thirty-nine received remedial instruction. (Again the purpose of the Mobile Reading Clinic was to determine pupils' reading deficiencies through diagnostic procedures). After proper identification of the deficiencies, the reading teachers were to assist boys and girls with their specific problems. These services were available to students of private and parochial schools; however, in the six months of operation no pupils from these sources participated in the Mobile Reading Clinic Project. To operate the Mobile Reading Clinic, a project manager was hired, with the responsibility of the overall supervision of the Mobile Reading Clinic. He acted as a liaison between the Mobile Center and the county, district, and school personnel. Fifteen reading teachers, who were certified or involved in the process of being certified in reading as well as having had classroom experience, were employed to work with the boys and girls. One of the requisites was a demonstrated willingness and enthusiasm to work with students in the disadvantaged areas. One secretary was employed to handle the necessary clerical operation.

The major role of the Mobile Reading Clinic was based upon the fact that approximately three to ten percent of the elementary school children in educationally deprived areas in Dade County were in urgent need of clinical diagnoses and instruction with their reading problem. These children had not succeeded in learning to read through existing programs nor had they been able to avail themselves of the services of the County district reading centers because of parental obligation to transport the child to and from the clinic. The cost of private services were an economic impossibility for the girls and boys in educationally deprived areas. Sixty-one elementary schools were designated as serving an educationally deprived population. Since the Mobile Reading Clinic was able to serve only ten of these schools, it readily became apparent that a number of students in need of the reading services were not being served.

The average time spent in the Reading Center for each child varied, depending upon the nature of the child and the nature of his reading disability. The average number of sessions attended was about forty-five.

It is anticipated that next year the major financial obligation will be salaries for personnel, instructional supplies, and equipment. No large amounts of money will be necessary for capital outlay. This is true only if the project is maintained at the present level of force.

EVALUATION

Design

In the evaluation of the Mobile Reading Clinic Project, the paragraph meaning, word meaning, spelling, and language subtests of the Stanford Achievement test were administered to all project students on a pre-post basis in order to determine the mean change.

A small group of students in the Mobile Reading Clinic was also matched with a group of remedial readers who were from the same schools but not in the project. Mean changes for these groups were compared.

An Informal Reading Inventory and the Spache Diagnostic test were also administered to the students in the Mobile Reading Clinic on a pre-post basis and the mean change reported.

The mean number of days absent during the spring semester of 1965-66 was compared with the mean number of days absent during the spring semester of 1964-65 for all students in the project to determine the extent of change.

Ratings along a continuum from "poor" to "excellent" were made by three groups of school personnel to assist in the total evaluation of the Mobile Reading Clinic project. Group I consisted of principals, project managers, district supervisors, and district superintendents. Group II consisted of teachers not in the project, and Group III consisted of teachers in the project.

Results

The results of the evaluation are reported in Tables A-G and Figures 1-5 in the ensuing section.

STANFORD ACHIEVEMENT TEST RESULTS

TABLE A

Total Sample Group of Treated Students

	GRADE 3			GRADE 4				GRADE 5				GRADE 6			
	P.M.	W.M.	S.	P.M.	W.M.	S.	L.	P.M.	W.M.	S.	L.	P.M.	W.M.	S.	L.
Pre	36	36	38.3	32.5	34.5	33.5	29.5	36	31.2	36.6	36.6	31.2	29.5	32.5	32.5
Post	35.2	32.5	37.2	32.5	29.5	37.2	31.2	36	32.5	36.0	37.2	32.5	29.5	32.5	32.5
Change	-.8	-3.5	-1.1	0.0	-5.0	3.7	1.7	0	1.3	-.6	0.6	1.3	0.0	0.0	0.0
N	39	39	39	28	28	27	24	34	34	39	34	40	40	39	39

MATCHED SAMPLE GROUPS - EXPEPRIMENTAL VS. CONTROL

TABLE B

Experimental Group

	GRADE 3			GRADE 4				GRADE 5				GRADE 6			
	P.M.	W.M.	S.	P.M.	W.M.	S.	L.	P.M.	W.M.	S.	L.	P.M.	W.M.	S.	L.
Pre	37.2	36.0	38.3	34.5	36.6	34.5	33.5	37.2	36.6	37.7	37.2	32.5	29.5	35.2	33.5
Post	34.5	32.5	37.2	32.5	31.2	38.3	32.5	36.6	32.5	36.0	36.6	33.5	31.2	35.2	35.2
Change	-2.7	-3.5	-1.1	-2.0	-5.4	3.8	-1.0	-0.6	-4.1	-1.7	-0.6	1.0	1.7	0.0	1.7
N	14	14	14	17	17	16	13	20	20	20	20	18	18	18	18

TABLE C

Control Group

	GRADE 3			GRADE 4				GRADE 5				GRADE 6			
	P.M.	W.M.	S.	P.M.	W.M.	S.	L.	P.M.	W.M.	S.	L.	P.M.	W.M.	S.	L.
Pre	37.2	36.0	39.2	35.2	36.6	37.2	36.0	37.2	36.6	39.2	35.2	32.5	31.2	33.5	31.2
Post	36.6	35.2	38.7	33.5	34.5	37.2	32.5	34.5	32.5	36.6	36.0	32.5	31.2	36.0	29.5
Change	-0.6	-0.8	-0.5	-1.7	-2.1	0.0	-3.5	-2.7	-4.1	-2.6	0.8	0.0	0.0	2.5	-1.7
N	14	14	14	17	17	15	12	20	20	20	20	18	18	18	18

TABLE D

Matched Groups Compared

		Experimental	Control	Difference
Mean change in Paragraph Meaning	(All Grades)	-1.0	-1.3	.3
Mean change in Word Meaning	(All Grades)	-3.7	-1.9	-1.8
Mean change in Spelling	(All Grades)	.2	-.2	.4
Mean change in Language	(All Grades)	.1	-1.1	1.2

TABLE E

Spache - Diagnostic Reading Scales Results

GRADE	N	MEAN SCORE PRE	MEAN SCORE POST	DIFFERENCE	t
3	21	1.2	2.0	.8	*5.97
4	20	1.9	2.6	.7	*6.14
5	31	3.0	3.8	.8	*9.64
6	42	2.8	3.5	.7	*12.96

TABLE F

Informal Reading Inventory Results

GRADE	Instructional Level					Independent Level					Frustration Level				
	N	Mean Score PRE	Mean Score POST	DIFF.	t	N	Mean Score PRE	Mean Score POST	DIFF.	t	N	Mean Score PRE	Mean Score POST	DIFF.	t
3	41	.7	1.5	.8	*9.64	38	.5	1.1	.6	*7.79	41	1.2	2.3	1.1	*13.25
4	38	1.3	2.1	.8	*7.02	37	.9	1.9	1.0	*9.62	37	1.9	3.1	1.2	*9.84
5	42	2.1	3.3	1.2	*12.77	42	2.1	3.2	1.1	*8.73	42	3.1	4.3	1.2	*11.01
6	12	2.0	3.2	1.2	**4.05	12	1.8	2.8	1.0	**3.47	12	3.0	4.1	1.1	**3.38

* Beyond .001 level of significance
** Beyond .01 level of significance

TABLE G

Attendance

CRADE	N	MEAN NO. DAYS ABSENT SPRING - 65	SPRING - 66	DIFF.*
3	38	6.1	5.2	-.9
4	41	7.5	8.0	.5
5	44	6.2	6.7	.5
6	44	5.8	5.4	-.4
TOTAL	167	6.4	6.3	-.1

*A negative difference indicates improvement since we are comparing mean number of days absent.

FIGURE 1

Ratings By Administrators

	Poor										Excellent
The Rating Scale	0	1	2	3	4	5	6	7	8	9	10
Frequency of Rating	0	1	0	0	1	1	4	1	3	4	2

N = 17   Mean 7.1   Bi-Modal 6,9   Median 8

FIGURE 2

Ratings By Teachers Not In The Project

	Poor										Excellent
The Rating Scale	0	1	2	3	4	5	6	7	8	9	10
Frequency of Ratings	17	17	10	8	25	11	10	16	18	31	29

Mean 5.73 Mode 9 Median 6

FIGURE 3

Graph Of Ratings By Teachers

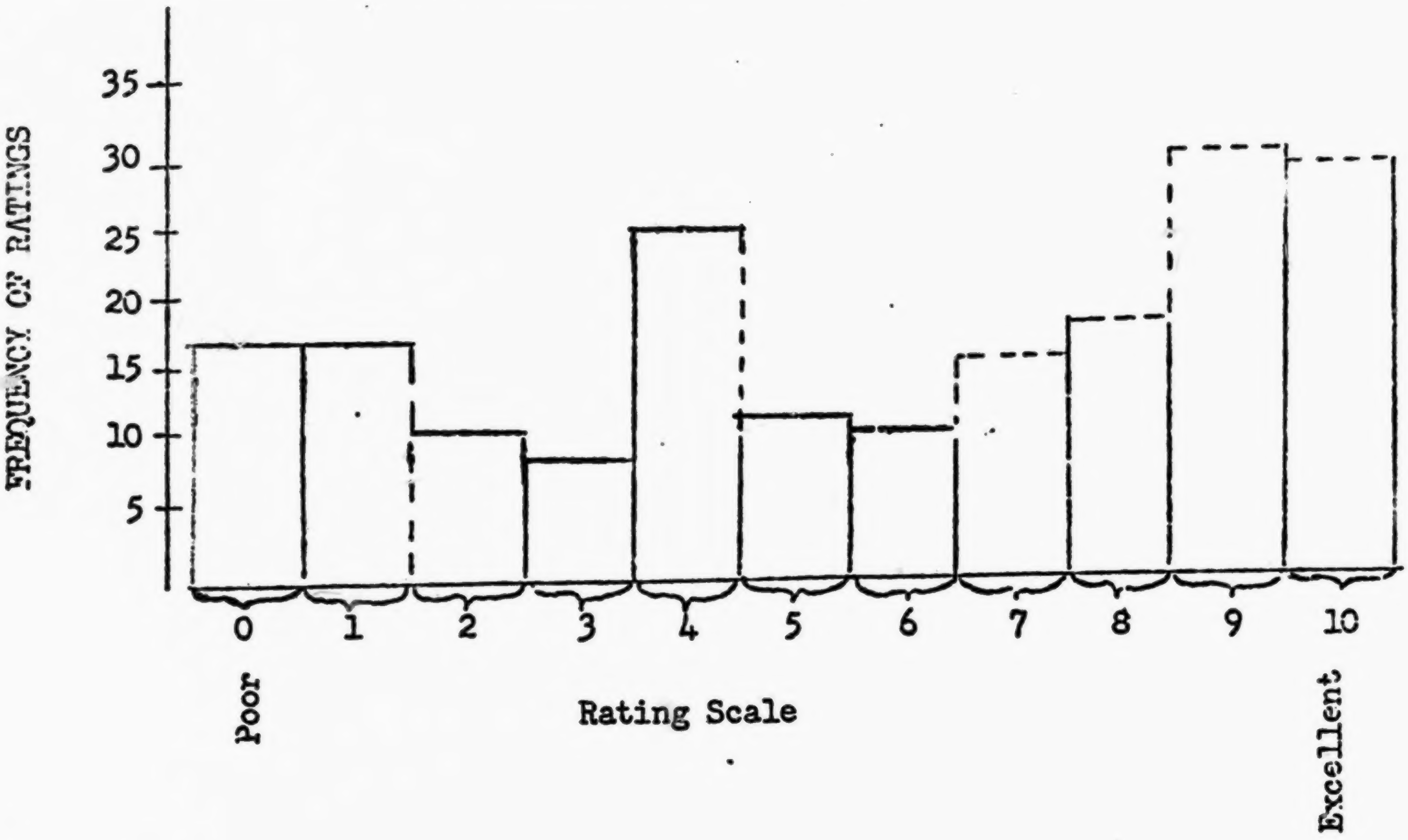


FIGURE 4

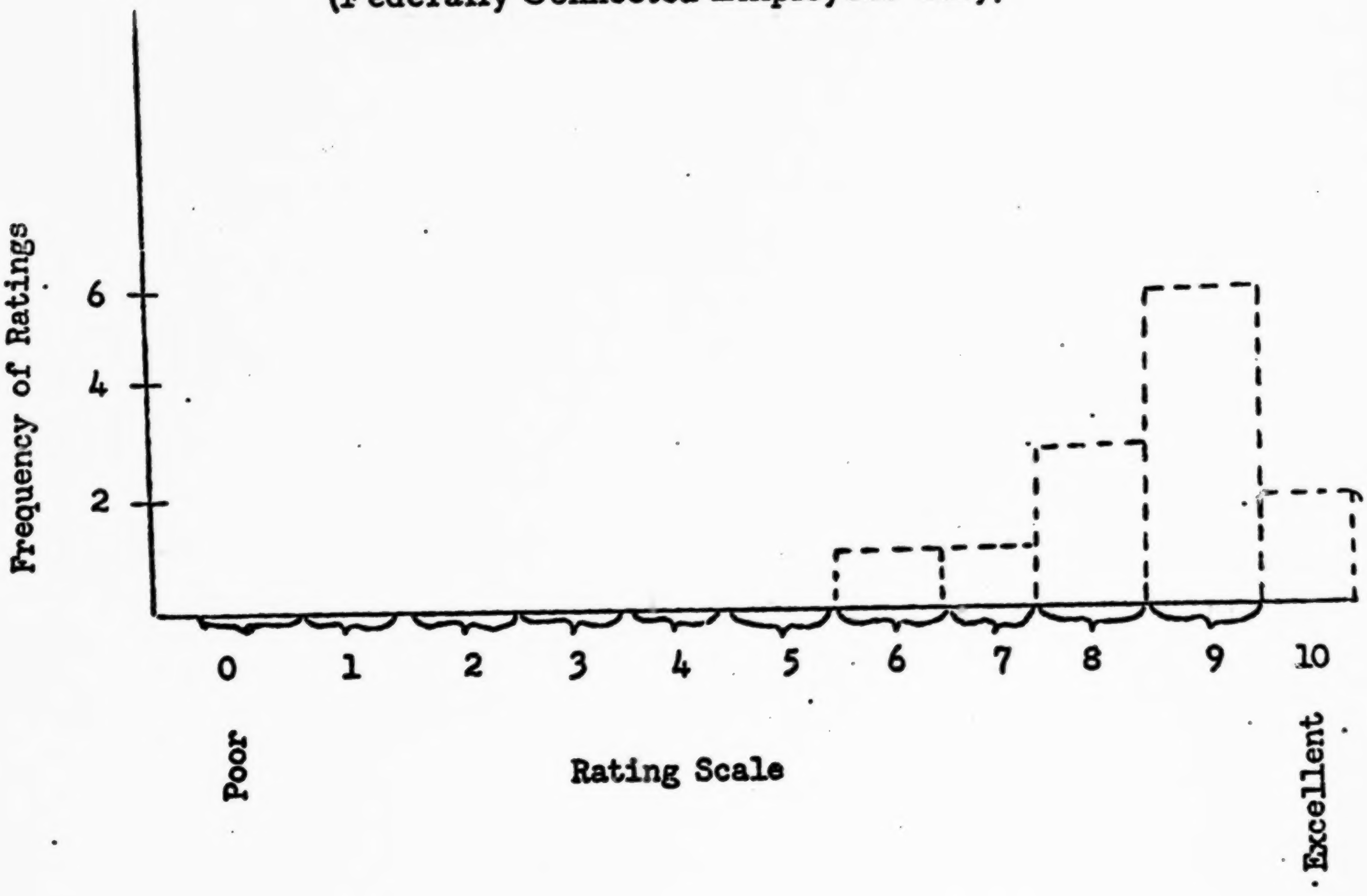
Ratings By Teachers In The Project

	- Poor										- Exc.
The Rating Scale	0	1	2	3	4	5	6	7	8	9	10
Frequency of Ratings	0	0	0	0	0	0	1	1	3	6	2

Mean 8.54   Mode 9   Median 9

FIGURE 5

Graph Of Mobile Reading Teachers Survey
(Federally Connected Employees Only)



Conclusions

The Mobile Reading Clinic project treated 239 pupils in grades three through six.

Significant positive mean gains were shown in the results from pre-post testing, using the Informal Reading Inventory and the Spache Reading Scales.

Small positive mean gains were also shown by the total Mobile Reading Clinic group in the Stanford Achievement subtests in paragraph meaning, spelling, and language on a pre-post test basis. A small experimental group evidenced a slight positive gain over the matched control group in paragraph meaning, spelling, and language.

In comparing the attendance of students in the Mobile Reading Clinic to their attendance during the same period of the previous year, it was concluded that attendance did not improve significantly.

The Mobile Reading Clinic received a very high rating on the survey of administrators' and teachers' opinions regarding the project.

COMMENTS RELATED TO PROJECTS

During June 1966, all teachers and administrators in schools receiving the services of a Title I of Public Law 89-10 project were permitted to comment concerning the program. Since comments were not required of respondents, it should be pointed out that the remarks listed are those of persons electing the option.

The comments of teachers pertaining to the Mobile Reading Clinic are listed below in the order of the frequency of number of similar responses.

> The supply of materials and equipment should be made available within a reasonable amount of time after it is ordered. Need for expanded equipment such as a Keystone Telebinocular exists.
>
> The program is doing an effective and important service.
>
> A medical unit should be employed, as many reading problems found have a health factor as a cause.
>
> In-service workshops with all mobile reading personnel should be scheduled.
>
> The program has tremendously increased student interest in reading.
>
> No more than two children should be served by each member of the unit at a time.
>
> More unified diagnosis should be used.
>
> More emphasis should be placed on teaching and less on written reports.
>
> Better means of communication should be formulated between the project manager and the units.
>
> Scheduled meetings should be held between classroom teachers and mobile reading teachers.
>
> There should be more mobile reading units, and they should be better equipped.
>
> The units should spend more time at an individual school, increasing the amount of treatment administered.

This program is excellent for increasing the vocabulary of students.

The regular classroom teachers need to be made aware of the total program.

The mobile reading unit does not remain at the schools long enough to be as effective as it could be.

There is a lack of communication between teachers and mobile reading personnel.

Reduced class loads would be most helpful for improving reading.

It is believed that this program is an excellent one and should be continued.

Case Study III

EVALUATION OF THE CULTURAL ENRICHMENT PROJECT

DETROIT PUBLIC SCHOOLS

Title: Cultural Enrichment Project

Funding: $996,070 under ESEA, Title I

Period: November 11, 1965 through August 31, 1966

Objective: To provide culturally enriching experiences to the 216,000 pupils attending public and parochial schools in disadvantaged areas of Detroit.

Procedures: Programs of cultural value were brought to the pupils in their schools, or pupils were transported to such events as could not be brought to the schools. After-school classes were conducted in the Fine Arts. A series of humanities-oriented programs in literature were broadcast, during school hours, over the educational broadcasting system of the public schools.

Evaluative Design: The original evaluation design listed six objectives in terms of pupil behavior and the types of evidence of attainments of the objectives. Briefly stated the objectives were that pupils would participate in a wide variety of cultural activities, that they would engage in creative activities, that the experiences would have both entertainment and educational values. Expected behavioral changes to be assessed were changes in reading and television viewing practices, in attendance at museums, art galleries, libraries, and the legitimate theater, and in participation in creative productions. Another expected outcome was the development of greater appreciation of the fine arts which would extend into the pupils' adult lives. The plan for the process observation called for observations and consultation to improve this program and its management, so that larger numbers of pupils could receive greater benefits from the project.

The Evaluation: The evaluators of the project were a research associate, already having full-time responsibilities and assignments in the research department, and a research assistant, assigned part-time to the project, beginning in June 1966. The associate had had considerable experience in public school research, instrument formulation, and report writing. The assistant was more highly trained in a doctoral program in educational research, research design, statistical techniques, and computer utilization, but had little practical experience in large scale research in a public school setting.

Limitations in the Evaluation: Because so little time could be given to the evaluation of this project, no attempt was made to measure changes in pupils' behaviors, attitudes, and appreciations. The evaluation of the program events facet of the project was limited to teachers' appraisals of the values of each program event and simple counts of numbers of pupils participating in each event. An estimate of the cost of each program event was made in terms of cost per participating pupil. A similar type of analysis was planned for the after-school enrichment classes. Because the television program was still in the developmental stage, no attempt was made to evaluate this phase of the project.

Findings: Product Evaluation. About 93,000 public school and 7,000 parochial school pupils participated in about 50 different program events including artist demonstrations, musical concerts, dramatic presentations by professional actors, and trips to cultural centers. Teachers rated all but three of the events as having high to very high interest and educational values. Costs ranged from an estimated $6.75 per pupil to a low of $.07. Values as reported by teachers were not necessarily proportionate to per-pupil costs.

EVALUATION OF THE PROGRAM EVENTS FACET OF THE CULTURAL ENRICHMENT PROJECT*

The Problem of Cultural Deprivation

Typical of the contrasts found within the inner city of Detroit are those of affluence and poverty and of culture and cultural deprivation. The center-city area of about 16 square miles encompasses some of the most valuable property in the city and the most dilapidated slums. The area contains the cultural center of the city--a state university, museums, the main branch of the public library, concert halls and auditoriums, and legitimate theaters--and a population that is relatively untouched by the cultural offerings. The poor cannot or do not take great advantage of the opportunities at their doorsteps. In the midst of plenty, there is cultural, as well as economic, poverty. One of the great challenges of education in the inner city is that of providing the opportunities and creating the desire for children to have culturally enriching experiences.

The Cultural Enrichment Project

The Cultural Enrichment Project was conceived as a program to bring to the 73,000 pupils at 59 public schools in the most economically deprived school communities a variety of experiences in the Fine Arts, and to bring to the 36,000 parochial school pupils and to the 126,000 pupils in 119 other public schools in areas of economic deprivation some of the culturally enriching experiences more readily available to children in neighborhoods of higher socio-economic status. The program, as conceived and carried out, brought to the children in their school settings a series of program events, including musical concerts, artist demonstrations, and plays by professional theater troupes. Also included under the program events facet of the project were the busing of pupils to places of cultural interest and to programs which could be presented better elsewhere than in the schools, and the provisions of free tickets to programs in concert halls. Other facets of the project, not included in this evaluative report, consisted of the production and presentation of a series of television programs on literature and the conduct of a large number of after-school classes which provided for children's active participation in the arts.

The Operation of the Project

The Cultural Enrichment Program was funded for the school year 1965-66 and for the summer of 1966 through a grant under the Elementary and Secondary School Act, Title I. A directing committee was set up to establish policies, select programs and artists, and to coordinate the program in the public and parochial schools. Staff, including a coordinator charged with the administration of the program, was obtained chiefly through the reassignment of Detroit Public Schools personnel. Contracts for performances were made with artists and organizations which were to produce the individual program events. The Research and Development Department of the Detroit Public Schools was assigned the responsibility of evaluating both the processes by which the program was carried out and the product of the program.

---

*Funded under the Elementary and Secondary Education Act, Title I

The remainder of this report deals with the evaluation of that facet of the Cultural Enrichment Program which provided the individual program events offered to the pupils.

The Evaluation Plan

The evaluation of a multi-faceted project such as the Program Events portion of the Cultural Enrichment Project (CEP), which involves pupils' exposure to and participation in culturally enriching experiences, cannot be based on a rigid research model. The evaluation of this project is, rather, based on examination of three criterion variables: (1) the number and variety of the experiences made available to pupils, (2) the numbers of pupils exposed to these experiences, and (3) the cultural values of the experiences. The discussion which follows deals with the extent to which the CEP meets its general objective of providing pupils in culturally, socially, and economically deprived areas with experiences which are culturally enriching.

Kinds of Program Events Provided

Experiences provided in the Program Events portion of the CEP fall into five general categories:

1. Musical events presented in the schools by individuals and by groups of performers, and musical events presented in concert halls for which free tickets were provided to pupils.

2. Artist demonstrations in which a commentator explained the processes by which a painter or sculptor created a picture or clay model in front of an audience.

3. Dramatic events, including plays by professional actors.

4. Trips to art institutes, museums, educational institutions, and zoos.
5. "Conversation in the Arts," after-school classes, offered to selected pupils at centralized locations.

Some 50 different programs of these types were presented. The list of the different program events is shown in Table I of this report.

The Number of Pupils Participating

Table I identifies the more than 50 different events and shows the numbers of public school pupils at each school level and the numbers of parochial school pupils at all grade levels who participated in each of the events. The totals from that table, along with the percentages of the CEP area school populations which participated in the program events, are as follows:

School Level	No. of Pupils Participating	School Enrollment	Percentage of Pupils Participating
Elementary	33,709	119,400	28
Junior High	27,435	31,500	87
Senior High	32,175	29,800	108
Parochial	7,088	35,700	20
Total	100,407	216,400	46

Product Evaluation

Values of the Program Events

There are two ways of evaluating the quality of each program. One is through inspection of the content of the program itself and of the quality of the artist presenting it. This type of evaluation was made by the CEP Advisory Committee prior to contracting for the services of the artists.

The other method of obtaining evaluation of program quality is through the appraisals of qualified observers at each program event. For the purpose of obtaining data for this type of evaluation, each teacher who accompanied a group to an event was asked to rate the event on each of four criteria: educational value, entertainment value, pupils' interest during the program, and the strength of his recommendation that the same program be repeated with other groups of pupils. A five-point rating scale was used, with "1" representing the lowest or negative value, and "5" the highest positive value. The mean of the ratings given each criterion by all the teachers viewing each program and a composite mean of the ratings given all four criteria were computed. The means of these ratings and the numbers of schools and pupils involved in each event are shown inTable I.

Observations and Recommendations Based on Findings

1. The program events have provided public secondary school pupils with an average of about one exposure per pupil to a culturally enriching event. The exposure of elementary and parochial school pupils is about one quarter that of the public secondary school pupils. Only 44 percent of the pupil population in the Project Area schools has had exposure to these programs. Even this percentage figure is misleadingly high because it is based on multiple counts of some pupils who may have had from two to seven exposures to program events.

   On the basis of these observations it is recommended that more pupil participation, especially for public elementary school and parochial school pupils, be provided by

   a. Offering more programs suited to elementary school pupils
   b. Establishing better communications between public and parochial schools with respect to attendance at the programs
   c. Persuading more principals to participate in the program offerings
   d. Arranging for busing pupils from schools having inadequate facilities to central locations where the different programs may be presented
   e. Scheduling at least one program event for each of the Project Area schools.

2. The impact of the remark "Look, they've got one of us 'cats' for conductor" and the wild applause accompanying the entry of Henry Lewis as conductor of the Detroit Symphony Orchestra was probably of more significance than the 4.9 rating given by 26 teachers. A less-than superior artist who has risen from the cultural milieu of the audience can be more effective than the superior artist from a different background. The cultural effect of Josephine Love on an audience of deprived youth may be greater than that of the Shaw Chorale. Idealistically, each performance should be judged on its own merits; practically, empathy may be as important as or more important than artistry.

3. Teachers' appraisals of the different events were very favorable. The

TABLE I

Means of Ratings Given by Teachers and Numbers of Pupils and Schools Participating in Cultural Enrichment Project Program Events

Event	Means of Teachers' Ratings*					No. of Schools	Number of Pupils				
	Educ. Value	Enter. Value	Pupil Interest	Recom. Repeat	Com-posite		El.	JHS	SHS	Paroch	Total
Musical Events											
Ashby Trio	4.5	5.0	4.4	5.0	4.7	8	1702	223	1800	590	4315
Brass Ensemble	5.0	4.9	4.9	5.0	4.9	17	3696	555	850	41	5142
K. Britten	4.3	4.4	4.2	4.4	4.3	9	335	1688	550		2573
R. Clark	3.0	3.0	3.0	3.0	3.0	1			400		400
Det. Sym. Orch.	4.9	4.9	4.8	5.0	4.9	26	4478	3295	5342	834	13949
Kallas Trio	3.0	5.0	5.0	3.5	4.1	1	315				315
J. Love	4.3	4.0	3.9	4.2	4.1	19	2692	771			3463
String Ensemble	4.9	4.6	4.7	5.0	4.8	30	3128	1035		2342	6505
Woodwind Ens.	4.7	4.4	4.5	4.8	4.6	25	3718	1787	250	947	6702
Total						136	20064	9354	9192	4754	43364
Concerts											
Det. Sym. Orch.	5.0	5.0	4.9	4.8	4.9	7		13	44		57
Folk Dance Fest.	4.6	4.9	5.0	5.0	4.9	5	332		25		357
All-City Honors	5.0	4.0	4.5	5.0	4.6	2			169		169
Jewell Chorale	4.8	3.8	4.0	4.5	4.3	5		2	94		96
NEHS Concert	5.0	5.0	5.0	5.0	5.0	1			870		870
HS Choral Fest.	5.0	5.0	5.0	5.0	5.0	1			54		54
Shaw Chorale	5.0	5.0	4.7	4.7	4.8	3			24	25	49
J. Walters	4.6	4.2	4.0	4.8	4.4	10		26	39		65
W. Warfield	4.8	4.7	4.5	5.0	4.7	6			62		62
Total						36	332	41	1381	25	1779
Artist Demonstrations											
Bostic	5.0	4.0	4.0	5.0	4.5	1			120		120
Brackett	4.7	4.6	4.6	4.9	4.6	7		905	400	394	1699
Gillerman	4.2	3.5	4.2	4.8	4.2	4		1399	450		1849
Graves	4.5	4.2	4.6	4.8	4.5	9	596	415		46	1057
McGee	5.0	4.6	5.0	5.0	4.9	5		510	700		1210
Mindener	5.0	4.6	5.0	5.0	4.9	5		432	386		818
Owens	3.0	3.0	3.0	3.0	3.0	1		400			400
Quinlan	4.5	3.5	4.5	5.0	4.6	2		177		250	427
Wald	5.0	3.7	4.7	5.0	4.6	3			620	300	920
Total						37	596	4238	2676	990	8500

(Continued)

TABLE I (continued)

Means of Ratings Given by Teachers and Numbers of Pupils and Schools Participating in Cultural Enrichment Project Program Events

Event	Means of Teachers' Ratings*					No. of Schools	Number of Pupils				
	Educ. Value	Enter. Value	Pupil Interest	Recom. Repeat	Com-posite		El.	JHS	SHS	Paroch	Total
Drama											
Just-So Stories	4.4	4.7	4.8	4.8	4.7	105	7969			392	8361
Man Who Married	4.3	4.8	4.8	4.8	4.7	31		12686	12273	85	25044
Old Maid & Thief	4.7	5.0	4.7	4.9	4.8	7			6054		6054
Proctor's Puppets	4.8	5.0	4.9	5.0	4.9	25	3204	500		739	4443
Young Tom Edison	4.6	4.9	4.6	4.7	4.7	7	313				313
Total						175	11486	13186	18327	1216	44215
Trips											
Art Institute	5.0	5.0	5.0	5.0	5.0	2	30	24			54
Child. Museum	5.0	3.0	5.0	5.0	4.5	1		23			23
Cranbrook	4.8	4.3	4.8	4.9	4.7	8	141	112		30	283
Greenfield Vill.	4.7	3.8	4.7	4.8	4.5	6	162	68	115		345
Historical Mus.	4.8	4.0	4.6	4.8	4.6	5	121	73			194
WSU Comm.Art Cent	4.6	3.5	4.1	4.2	4.1	18	326	204			530
Zoo	5.0	5.0	5.0	5.0	5.0	3	311				311
Miscellaneous	4.6	3.5	4.4	4.8	4.3	7	100	112	130	49	391
"Were You There"	4.8	4.2	4.2	4.6	4.6	5	40		73		113
Total						55	1231	616	318	79	2244
"Conversations"											
French Art	5.0	5.0	5.0	5.0	5.0	1			33		33
Mich. History	5.0	5.0	5.0	5.0	5.0	1			90		90
Folk Art	5.0	5.0	5.0	5.0	5.0	1			38		38
Art Apprec.	5.0	5.0	5.0	5.0	5.0	1			35		35
Books	5.0	3.0	5.0	5.0	4.5	2			29	8	37
Sculp.Paint,Des.	5.0	5.0	5.0	5.0	5.0	2			32	12	44
Literature	5.0	3.0	4.0	5.0	4.2	2			18	4	22
Symph. Music	4.0	3.0	4.0		3.7	1			6		6
Total						11			281	24	305
Grand Total							33709	27435	32175	7088	100407

*Ratings are made on a 5-point scale, with "1" being "very low" and "5," "very high" value.

September, 1966

composite ratings of the values of the events were, with the exception of three events, between "high" and "very high." The most highly rated programs, for which an adequate number of appraisals were received, were those with a composite rating of "4.9." These are the following:

The Brass Ensemble from the Detroit Symphony Orchestra
The Detroit Symphony Orchestra
    (presentations in the schools or at the Ford Auditorium)
The Detroit Public Schools Folk Dance Festival
Artist demonstrations by McGee and by Mindener
Proctor's Puppets

Following closely in value, as represented by teachers' composite ratings of "4.8" and 4.7," are the following:

The Ashby Trio
The String Ensembles from the Detroit Symphony Orchestra
The Shaw Chorale, an outside concert
The W. Warfield Concert, an outside concert
The ballet presentations of Kipling's "Just-So Stories"
The Milan Theater Group's Presentations of "The Man Who
    Married a Dumb Wife"
The light operetta "The Old Maid and the Thief"
The light operetta "Young Tom Edison"
Visits to Cranbrook Academy

It is recommended that, in the absence of other contra-indicative data, these programs receive high priority in the plans for the CEP for the coming year.

4. The following programs given ratings from "4.3" through "4.6" were appraised by the teachers as having "high" educational and interest values:

K. Britten
Woodwind Ensemble
All City Honors Concert
J. Walters Concert
Artist demonstrations by
    Bostic
    Brackett
    Graves
    Quinlan

- Wald
- Trips to
  - Children's Museum
  - Greenfield Village
  - Historical Museum
  - Program "Were You There"
- "Conversations" program on books

5. The numbers of reports made on some other program events are inadequate to supply valid appraisals of these programs. This observation is especially true of any program rated "5.0" by a single observer, in some cases the sponsor of the program he rated. Among such programs are the following:

   - The Northeastern High School Concert
   - The High School Choral Festival
   - Trips to the Art Museum and the Zoo
   - "Conversations" classes in
     - French Art
     - Michigan History
     - Folk Art
     - Art Appreciation
     - Sculpture, Painting, and Design

   Of equally doubtful value are the ratings on three program events receiving ratings between "3.0" and "4.0." Whether these comparatively poor ratings were due to the fact that the artist lacked experience before school-aged audiences (each of these programs was offered only once), to the intrinsic lack of merit of the program itself, or to a bias of the rating teacher is not known. The following programs received the lowest ratings:

   - Musical programs by Clark
   - Artist demonstration by Owens
   - "Conversations" class in "Appreciating Symphonic Music"

On the basis of the above mentioned facts, it is recommended that further evaluative data be obtained before a decision be made regarding the retention of these programs in the CEP. It is, however, the opinion of the evaluator that the two school-produced programs, the Northeastern High School Concert and the High School Choral Festival, offer such unique opportunities for both active and passive participation by pupils that they, or similar programs, should be retained in the project, and that more programs of this type should be offered.

6. Certain programs received composite ratings of "4.0" through "4.2." While these are "high" ratings, they are relatively low in comparison with those given other program events. These programs are the following:

   The Kallas Trio
   Recitals of Josephine Love
   Artist demonstration by Gillerman

   Visits to the Wayne State University Community Arts Center
   "Conversations" class in "Literature"
   Miscellaneous trips

   The continuance of these programs is given qualified recommendation, with the suggestion that their strengths and weaknesses be studied with the purpose of increasing their value to pupils.

7. Other recommendations, based on an analysis of the program offerings rather than on the evaluative data, are the following:

   a. That inner-city school children have the opportunity to find meaningful cultural enrichment through trips to farms having domestic animals
   b. That the CEP program be expanded to include more program events of a literary nature. The "Poetry Circuit" program planned for the year 1966-67 is a step in the recommended direction.
   c. That CEP funds be used to set up exhibits of pupil work produced under the aegis of the CEP after-school program. (A combination of an arts and crafts exhibit with musical, dance, and/or dramatic programs presented by inner-city pupils could serve to raise self concepts, show the public some of the products of ESEA Title I funds, and provide enriching experiences for both pupils and adults.)

8. It might be noted that a study of the effectiveness of another project, the Tri-Area Integration Project, involving about 50 schools in the Mackenzie, Mumford, and Pershing constellations, investigated the effectiveness of various extra services in meeting the objectives of the project. The responses of teachers and administrators in these schools rated the teacher aides and cultural enrichment experiences as being the greatest contributors to the objectives of maintaining the integrated character of the neighborhood,

supplying high quality education, providing additional learning experiences, and developing harmonious school-community relationships. In the same study over 50 percent of presidents of school organizations and of lay members of the Tri-Area Integration Project school committees said that the CEP was of great or very great importance in providing educational opportunities for pupils.

Process Evaluation

The preceding account of teachers' appraisals of individual program events sponsored by the Cultural Enrichment Project attests to both the magnitude and the excellence of this phase of the project. The paucity of negative criticism on the report forms shows that the programs were well received by both staff members and pupils. Both the selection of the program events and the administering of this phase of the program has, on the whole, been successful, as is shown in Table I.

There are, however, certain additional data contained in Tables II and III of this report that should be carefully scrutinized before plans are made for next year. The Process Evaluation which follows deals with three phases: (1) the scope of evaluation, (2) the distribution of program events, and (3) the participation by parochial schools. Suggestions are made, even though the data on which they are based are admittedly incomplete.

Scope of Evaluation

The scope of evaluation is limited in that it does not get reactions directly from pupils. Teachers' judgments and impressions are valuable, but they are not necessarily a reflection of pupils' values. Furthermore, the evaluations and reports of pupil attendance are incomplete. Reports have been received from about 90 percent of the schools participating in all individual program events. While this is a high percentage of returns, some method should be derived for obtaining reports from all teachers accompanying pupils to program events. The appraisal forms must be revised to include a statement of numbers of pupils from other public and parochial schools, as well as the numbers from the host school.

Continuing evaluation was such that, while the programs were being presented, teachers' criticisms and suggestions for improvement of individual program events were transmitted to the supervisors of the different types of events. Subsequent conferences with the artists resulted in improvements in their programs. Likewise, the evaluator's recommendations to the project coordinator and to the directing committee resulted in some changes in operational procedures and in amelioration of some of the inequities of distribution of programs.

Distribution of Program Events in Public Schools

Inspection of Tables II and III reveals different types of difficulties in distribution of program events. Among the 137 elementary, 24 junior high, and 15 senior high schools listed as being project schools, 48 elementary, 4 junior high, and 3 senior high schools did not report any participation in major program events; 45, 3, and 2 schools in these categories did not report any participation whatsoever by their pupils. On the other hand, 15 elementary, 11 junior high, and 8 senior high schools reported participation in 4 to 7 major events. In terms of ratios of participating pupils to school enrollments, the elementary schools' ratio was under .21 for 83 schools and over .80 for 12 schools; among the junior high schools, the ratio was under .21 for 4 schools and over .80 for 10 schools; among the senior high schools, the ratio was under .21 for 4 schools and over .80 for 5 schools.

Table II

Frequencies with Which Schools Reported Participation of Thirty or More Pupils in Program Events of the Cultural Enrichment Project

No. of	Number of Schools Reporting Participation				
Events	Elem.	Junior High	Senior High	Parochial	Total
0	48	4	3		55
1	41	3	1	20	65
2	17	3	2	1	23
3	16	3	1	3	23
4	11	2	3	1	17
5	1	7	3		11
6	3	2			5
7			2	1	3
Total	137	24	15	26	202

Table III

Ratios of Pupils Participating in Program Events of the Cultural Enrichment Project to Total School Enrollments*

Ratio Interval	Number of Schools Reporting Participation				
in Percents	Elem.	Junior High	Senior High	Parochial	Total
0 %	45	3	2		50
1-10	18		1		19
11-20	20	1	1		22
21-30	6	1			7
31-40	9	2			11
41-50	5		1		6
51-60	10	2			12
61-70	5	2	3		10
71-80	2	1			3
81-90	7		1		8
91-100	1	3	1		5
Over 100	4	7	3		14
Unknown	5	2	2	26	35
Total	137	24	15	26	202

*For the purpose of this analysis the total enrollment was that printed in Detroit Public Schools Directory, 1965-66.

These figures and other related information point out several planning needs:

1. Presentation of more program events for the younger pupils
2. Persuasion of more principals to request program events
3. Provision for pupils in schools which lack adequate facilities to participate in program events offered in buildings having better facilities

In conjunction with the fact that a large number of project shools participated to only a slight degree in program events, it might be noted that teachers' evaluation reports indicated participation by non-project schools in program events as follows:

Program Event	No. Of Schools	No of Pupils
K. Britten	1	155
String Ensemble	1	173
Artist Demonstration	1	177
"Just-So Stories"	24	1650
"Old Maid and the Thief"	1	1600
"Man Who Married a Dumb Wife"	11	7143
Proctor's Puppets	2	387

The principal cost of the program "Just-So Stories" was that of transporting pupils to the program. While 1650 pupils in non-project schools were bused to this program, the transportation cost of these pupils was not charged to the project. The play "The Man Who Married a Dumb Wife" was produced in all non-project schools to fill open production dates caused by the inadequate facilities in project schools.

Participation of Non-Public Schools

Tabulations of teacher evaluation reports indicate that about 7000 non-public school children attended individual program events, including trips, concerts outside school buildings, and performances given in either public or parochial school buildings. This figure is approximately 7 percent of the total number of public participants. Furthermore, among the 68 listed parochial project schools, only 19 schools reported pupil participation, although reports of pupil participation were received from six parochial schools which were not listed as project schools. Written comments emanating from some of the public schools indicate that, although 20 percent of the auditorium seating capacity had been reserved for non-public school pupils, the seats thus reserved were not used.

These facts elicit the following comments and suggestions:

1. There is need for a better method of communication, both in the issuance and acceptance of invitations, for parochial school pupils to attend events in public school buildings.

2. More program events might well be scheduled in parochial schools having adequate facilities for their presentation. While it is not practical to present a symphony orchestra or a play in a school without an auditorium, there are other program events which could be offered in large classrooms or gymnasiums.

3. A comparison of teacher appraisal reports received with the program events schedule for parochial schools reveals the fact that no appraisal and attendance report was received for seven of the scheduled events. As with the public schools, efforts should be made to obtain reports from all the adults accompanying pupils to program events.

4. If the non-public schools wish to participate in CEP program events, some of the responsibility of obtaining wider participation from more schools rests with the parochial schools themselves, as well as with the project coordinator.

Cost Analysis

Costs as Related to Benefits Reported

The report has shown thus far that teachers rated almost all of the program events as having high education, entertainment, and interest values and strongly recommended that they be repeated. These recommendations were made without consideration of the costs of the events. However, cost--both total and per pupil--should be considered in the evaluation of the program of the past year and in the planning for the present year. The discussion which follows deals with the cost of each program event.

Table IV gives, for each program event, the estimated numbers of participating pupils (allowing for participants not included in teacher appraisal reports), the total cost of the event (as reported by the Special Projects office and adjusted to include the estimated cost of busing project pupils), and the estimated cost per participating pupil. The costs included in this table do not include administrative and overhead costs for the entire CEP. The composite ratings given by teachers (Table I) are included in Table IV so that the estimated cost per pupil may be compared with the teachers' appraisal of the value of each program event.

TABLE IV

Cost and Appraisal Data on Individual Project Events of the Cultural Enrichment Project

Program Event	Composite Teachers' Rating	Estimated No. of Pupils Participating	Total Cost of Events	Estimated Cost Per Pupil
Musical Events				
Ashby Trio	4.7	4,745	$ 1,500	$ .32
Brass Ensemble	4.9	5,655	3,000	.53
Kay Britten	4.3	2,830	300	.11
R. Clark	3.0	440	30	.07
Detroit Symphony Orchestra	4.9	15,175	90,000	5.90
Kallas Trio	4.1	345	65	.19
Josephine Love	4.1	3,810	690	.18
String Ensemble	4.8	7,155	3,125	.44
Woodwind Ensemble	4.6	7,370	4,680	.63
Concerts				
Detroit Symphony Concert	4.9	120	180[a]	1.50
Folk Dance Festival	4.9	332	---[b]	
All City Honors Choir	4.6	29	---[b]	
Jewell Chorale	4.3	300	300[a]	1.00
NE High School Concert	5.0	870	---[b]	
High School Choral Festival	5.0	54	---[b]	
Shaw Chorale	4.8	150	300[a]	2.00
Walters Concert	3.8	100	150[a]	1.50
Warfield Concert		150	150[a]	1.00
Artist Demonstrations				
Bostic and Commentator	4.5	130	75	.57
Brackett and Commentator	4.6	1,870	450	.24
Gillerman and Commentator	4.2	2,035	300	.15
Graves and Commentator	4.5	1,165	550	.47
McGee and Commentator	4.9	1,330	375	.28
Mindener and Commentator	4.9	910	300	.33
Owen and Commentator	3.0	440	75	.17
Quinlan and Commentator	4.6	470	225	.48
Wald and Commentator	4.6	1,010	300	.30
Drama				
"Just-So Stories"	4.7	11,804	8,858[c]	.75
Milan Theater Productions	4.7	27,550	171,988[d]	6.25
"Old Maid and The Thief"	4.8	6,660	3,250	.49
Proctor's Puppets	4.9	4,890	425	.09
"Young Tom Edison"	4.7	345	493[c]	1.43

TABLE IV (Continued)

Cost and Appraisal Data on Individual Project Events of the Cultural Enrichment Project

Program Event	Composite Teachers' Ratings	Estimated No. of Pupils Participating	Total Cost of Events	Estimated Cost Per Pupil
Trips				
Art Instutute	5.0	60	$ 33[c]	$ .55
Children's Museum	4.5	25	35[c]	1.40
Cranbrook	4.7	310	400[c]	1.30
Greenfield Village	4.5	380	300[c]	.79
Historical Museum	4.6	215	175[c]	.81
WSU Community Arts Building	4.1	585	315[c]	.54
Detroit Zoological Garden	5.0	340	135[c]	.40
Miscellaneous	4.3	430	245[c]	.57
Conversations Programs (12 programs for 30 pupils each, 6 sessions each program)		360	10,000 (140 per session)	28.00 (4.68 per session)
French Art	5.0			
Michigan History	5.0			
Folk Art	5.0			
Art Appreciation	5.0			
Books	4.5			
Sculpture, Painting and Design	5.0			
Literature	4.2			
Symphonic Music	3.7			
Miscellaneous				
"Were You There"	4.6	175	87.50[a]	.50

[a] Accurate data on attendance at these programs are not available. Since some pupils did not use the tickets they were given, the per-pupil costs are higher than the reported estimated costs, which correspond to prices paid for tickets.

[b] These events were not charged to the Cultural Enrichment Project.

[c] Total cost includes estimated busing cost for transporting project school school pupils.

[d] Total cost does not include the cost of equipment to which the Board of Education retains title.

Costs as Related to Other Factors

Analyses of costs per pupil reveal that some of the costs are high--higher in some cases than the cost of sending pupils directly to plays and concerts presented in commercial theaters and concert halls. On the basis of cost alone, it would be advisable to purchase and distribute to the pupils free tickets to plays and concerts. There are, however, other considerations besides those of cost. Some of these considerations are listed below.

1. If the large majority of pupils is to be exposed to a culturally enriching experience, the experience must be brought to the pupils in school during the school day.
    a. Experience has shown that many free tickets are not used, even though the recipients are carefully selected and have requested the tickets.
    b. The lateness of the hour and the location of the presentation of plays and concerts makes it inadvisable to give tickets for evening performances to any but the older and more mature high school pupils.
    c. The voluntary aspect of using a free ticket weakens the whole CEP program. Many pupils who live in the inner city do not have the cultural perceptions which would make them want to give up their leisure occupations to go to a concert or play. These perceptions must be inculcated in the school.
    d. Tickets to symphonic concerts and to good plays are often at a premium. For example, almost all seats to concerts by the symphony orchestra are sold to season ticket holders and sponsors of the orchestra.
2. Concerts and plays offered for the general public are attuned to a more sophisticated audience than school pupils. The development of perceptions and appreciations must have explanation to accompany the event. Therein lies their educational value.
3. Certain types of programs, especially those for elementary school children, require an intimate setting. Questions and group discussions cannot be efficiently handled in a large auditorium.

Recommendations Related to Costs

With these and other considerations in mind, the following recommendations are made:

1. Since the pupil costs of the touring theater group ($6.25 per pupil) and the symphony orchestra ($5.90 per pupil) are high, the values of these programs should be reappraised in terms of the objectives of the project. If it is decided to retain them in the CEP, effort should be made to reduce the pupil cost by more efficient use of existing facilities. The following suggestions are made to accomplish this:

    a. The scheduling of the programs in the auditoriums of high schools, with pupils bused from their own schools. Under this plan no rental would be paid for the use of auditoriums, and pupils in nearby schools could walk from their schools.

    b. Establishing and abiding by a firm rule that schools invited to send pupils to view a program accept or reject the invitation within a specified period of time. If the invitation is rejected, another school could be invited to participate. If the invitation is accepted but pupils are not sent to the program, the defaulting school would be considered as having participated, thus reducing its quota of events.

2. The costs of the "Conversations" programs per pupil per session are high. Seventy-two sessions conducted at a cost of $10,000 means that each session costs about $140. (Public school teachers receive less than $6.00 per hour for teaching in evening schools, and consultants' fees are from $50 to $100 per day.) A reappraisal of this series of program events is recommended to determine its values in terms of pupil benefits. It is possible that equal benefits might be derived at a smaller cost by the use of competent staff members from the schools or the universities in Detroit.

Conclusion

The product evaluation of the CEP has shown that the directing committee and the viewing teachers consider that the program events offered by the CEP had high educational and entertainment values and that they held pupils' interests. Lacking in the product evaluation are data on pupils' reactions to the program events. (This weakness should be corrected in the evaluation of the 1966-67 project.)

The process evaluation shows the need for more accurate and complete reporting of pupil participation in each event, the need for better coordination of public and parochial school planning and attendance at program events, the need for more programs for elementary school children, and the need for more efficient scheduling of high cost programs (perhaps in large auditoriums in the schools) so that wider pupil participation may be obtained at relatively small increases in the total costs.

Considering the gargantuan task of coordinating the activities in a developing program, of dealing with performers, school people, the directing committee, and the coordinators of programs--the administration of the entire CEP has been remarkably efficient. There is, however, room for improvement in the quality of some of the program events and in the administration of the program as a whole. It is expected that with the experience gained in 1965-66, with more time for planning, and with improved coordination among the CEP and other ESEA, Title I, projects, the program will be improved considerably in 1966-67.

It still remains to be determined what long-range effects the program will have on pupils' attitudes and behaviors.

Case Study IV

EVALUATION OF THE PRESCHOOL PROGRAM

LOS ANGELES CITY SCHOOL DISTRICTS

Spring Semester, 1966

INTRODUCTION

The Elementary and Secondary Education Act projects in the Los Angeles City School Districts included eleven elementary, twenty-five secondary, six special education, one urban affairs, and four auxiliary services. Projects were divided into activities which were evaluated and reported on separately. In interpreting the evaluations, the reader should consider that the projects were operative for no more than five months.

Each activity report has a similar format, and each activity has a code designator assigned. The code designator relates the activity to instruments used in the evaluation.

The activity report format is outlined below:

- 1.00 Description
- 2.00 Objectives
- 3.00 Implementation
    - 3.10 Duration of Project and Number of Schools
    - 3.20 Pupils
    - 3.30 Non-Public School Pupils
    - 3.40 Activities
        - 3.41 Staff Activities
        - 3.42 Pupil Activities
    - 3.50 Specialized Materials, Supplies, and Equipment
    - 3.60 Personnel and Logistical Problems
- 4.00 Evaluation
    - 4.10 Design
    - 4.20 Attainment of Objectives
        - 4.21 First Objective
        - 4.22 Second Objective
        - 4.23 Third Objective
    - 4.30 Outcomes
- 5.00 Conclusions
- 6.00 Recommendations

Under section 3.00, Implementation, any sub-section not a part of the report is omitted, but the numbering sequence is retained. Under section 4.20, data relating to each objective are summarized and analyzed. The cycle is repeated to evaluate each design objective.

The detailed project design for each activity report is included in Addendum A. Supplemental data may be found in Addendum B.

Computer services were sub-contracted to the Computer Sciences Laboratory of the University of Southern California and performed on the Honeywell H-800. Statistical programs written by consultants of the Office of Research and Development, and the HILOW and MOWI programs furnished by the Computer Science Laboratory were used.

PRE-SCHOOL - II AND IV
ESEA - 19-520-02-0014 AND 19-520-0016

1.00 DESCRIPTION

The Project II Pre-School Program established a pre-school class in each of 14 schools. The pre-school component of Project IV provided 20 classes in 19 schools. The classes were limited to an enrollment of 15 pupils and were scheduled daily from 9:00 a.m. to 11:30 a.m. or from 12:30 p.m. to 3:00 p.m. The program was directed toward extending and enriching the experiential background of the pupils in order to enable them to better adjust to and succeed in school.

2.00 OBJECTIVES

-To increase children's speaking and listening skills prior to their entry into grade 1

-To help children learn to play and work together in a social situation.

-To teach children to care for themselves, their personal needs, and their possessions in a group setting

-To enhance the pupil's opportunity for immediate success upon entering grade 1

-To provide pupils and parents with supplemental services

-To provide teachers with special materials, equipment, and in-service education

3.00 IMPLEMENTATION

3.10 Duration of Project and Number of Schools

The project was in operation from January 31 through June 17, 1966, with programs conducted at 14 Project II and 19 Project IV schools.

3.20 Pupils

The Project II Pre-School Program enrolled 208 children, and the pre-school component of Project IV served 297 children. The children were selected for the program on the basis of applications by their parents.

An enrollment procedure similar to that required for kindergarten pupils was utilized and supplemented by a parent-teacher conference. Eligibility for enrollment was also limited to pupils who would have been eligible to enroll in kindergarten for the 1966 fall semester. The classes were established with 15 pupils each under the direction of a qualified teacher. Paid teacher aides and volunteers, including parents, were used to maintain a ratio of one adult to five children. A concerted effort was made to encourage the visitation and participation of parents.

3.40 Activities

3.41 Staff Activities

Six days of training for the teachers were scheduled prior to the implementation of the pre-school program in the schools. The pre-service series included an overview of the program, a consideration of the developmental characteristics of pre-school children and the implications for the program, suggestions for organization and planning, and information pertaining to the administration of the program. In-service education meetings were held monthly by the pre-school consultants to assist the teacher in implementing the program.

The afternoon or second half of the teacher's day was utilized to plan and to extend the classroom program. Activities included specific planning for the next day, home visitations, individual pupil and parent conferences, group meetings with parents, maintenance of anecdotal records and records of daily activities, and acquisition of supplies and materials.

3.42 Pupil Activities

The pre-school program was centered on pupil activities which encouraged the development of perceptual-motor skills, appropriate social-emotional behavior, and a readiness for successful intellectual-academic behavior. The unique experiences which were provided for the pupils included observing and caring for plants and animals, participating in dramatic representation in the playhouse center, manipulating puppets, using a toy telephone, wheel toys, and playground equipment, singing and listening to music, exploring art media, looking at books and listening to stories, and viewing films and listening to records and tape recordings.

3.50 Specialized Materials, Supplies, and Equipment

Both Project II and Project IV schools received an allotment of consumable and non-consumable supplies including such items as balls, tempera paint, construction paper, paste, crayons, scissors, puzzles, dolls, and records. The Project II classes also received basic equipment which included bookcases, chairs, tables, easels, tape recorders, and autoharps. The school library made available picture and story books.

3.60 Personnel and Logistical Problems

Personnel problems included securing replacements for teachers assigned to these projects, the lack of paid aides, and the difficulty in obtaining and retaining reliable volunteer aides. Logistical problems included the late arrival of supplies and equipment and the lack of adequate classroom facilities and playground space in some schools. The lack of paid or volunteer aides was resolved in many instances by the use of interested parent volunteers. The lack of classrooms was resolved by organizing pre-school classes on a morning and afternoon basis with two teachers sharing one room.

4.00 EVALUATION

4.10 Design

Project objectives were evaluated according to the following variables: standard pronunciation of words, oral expression of ideas, number of critical incidents, self-care and hygiene, development of school readiness, parent reactions, and staff evaluation of the overall effectiveness of the program.

The following instruments were employed to collect information on the variables:

-Form E4A appraised pupil behavior and progress.

-Form E4B assessed pupil's physical-motor development.

-Form E4C appraised administrative opinion of the program's effectiveness.

-Form E4D appraised teacher-consultant opinion of the program's effectiveness.

- Form E4E secured teacher opinions of the program's effectivness.
- Form E4F sampled parent reactions to the program.
- Peabody Picture Vocabulary Test assessed pupil readiness for school.
- Vineland Social Maturity Scale gathered data from parents.

4.20 Attainment of Objectives

4.21 Objective: To increase children's speaking and listening skills prior to their entry into grade 1.

4.22 Objective: To help children learn to play and work together in a social situation.

4.23 Objective: To teach children to care for themselves, their personal needs, and their possessions in a group setting.

Information in Table A is based upon a teacher rating scale which compares pupil growth in relation to the above objectives.

Teacher ratings indicated that pupils showed improvement in speaking and listening skills, in social and emotional development, and in personal-emotional development. Analysis of mean differences from pre-and post-tests was significant at the .01 level.

TABLE A

Mean Scores of Teacher Ratings

Item	Pre	Post
Child recognizes major body parts	3.5	4.1*
Child responds verbally to questions during conversation	3.2	3.7*
Child responds non-verbally to questions during conversation	2.5	2.3
Child pronounces words correctly	3.1	3.5*
Child asks questions which imply an understanding of what has been explained	2.7	3.3*
Child speaks freely and is able to verbalize a need	3.3	3.8*
Child uses words correctly in terms of their meanings	3.2	3.7*
Child is able to reach an amicable agreement with another child regarding shared use of equipment and materials	2.9	3.4*
Child takes active part in group activities	3.4	3.9*
Child is aware of space limitations, i.e. pouring liquid into vessels or loading blocks on a cart	3.2	3.9*
Child seeks out another activity following a peer conflict decided in favor of another child by the teacher	2.8	3.4*
Child willingly acts upon teacher requests	3.5	3.8*
Child initiates conversational exchanges with adults	2.9	3.4*
Child requires no assistance in restroom activities	2.5	2.6
Child utilizes alternative approach when initial method of problem solving proves insufficient	2.9	3.3*

*t-score significant at the .01 level

Table A is based on Form E4A

4.24 Objective: To enhance pupils' opportunities for immediate success upon entering grade 1.

The Peabody Picture Vocabulary Test was administered in a pre- and post-test situation in order to obtain intelligence test scores. Form E4B was also administered on a pre- and post-basis to gather data dealing with the psychomotor development of the pupils. The Vineland Social Maturity Scale was administered by counselors to the parents of the pupils in six randomly selected schools. Results of this scale were interpreted for teachers by counselors in order to assist them in meeting the individual needs of the pupils.

Results of the Peabody Picture Vocabulary Test and Psychomotor Development Test (Form E4B) indicated that pupils made improvement in school readiness. The mean of the pre Peabody Picture Vocabulary test was 36.7 and the mean of the post test was 47.5. Difference between the means was significant at the .01 level. The mean of the pre Psychomotor Test was 4.8 and the mean of the post test was 7.0. Analysis by t test indicated significance at the .01 level.

4.25 Objective: To provide pupils and parents with supplemental services.

Table B shows parent reactions to the effectiveness of the program. Ninety-five percent of the parents stated that their children learned "quite a bit" or "a great deal" in the pre-school class; 99 percent said that the program should be continued.

TABLE B

Parent Reaction To The Program

	Almost None		Very Little		A Fair Amount		Quite a Bit		A Great Deal	
	f	%	f	%	f	%	f	%	f	%
Did your child show interest in the program?	0	0	3	1	20	5	153	38	243	56
Did your child learn from the class?	0	0	0	0	21	5	120	30	279	65

	Yes		No		Not Sure	
	f	%	f	%	f	%
Should the class be continued?	419	99	1	1	0	0

Table B is based on Form E4F, items 1, 2, and 3. N=420

4.26 Objective: To provide teachers with special materials, equipment, and in-service education.

Table C shows how teachers rated the availability of supplies, equipment, audio-visual materials, and consultant time. Teacher comments indicated that supplies and equipment were not available at the onset of the program. Twenty-two of thirty-eight teachers indicated that equipment at their disposal was sufficient.

TABLE C

Teacher Evaluation

		Supplies		Equipment		Audio-visual Materials		Consultant Time	
Scale from 1 to 5		f	%	f	%	f	%	f	%
Insufficient	1	2	5	7	18	3	8	1	3
	2	6	16	9	25	7	18	2	5
	3	8	21	7	18	6	16	3	8
	4	10	26	5	12	6	16	8	21
Sufficient	5	12	32	10	26	16	42	24	63

Table C is based on Form E4E.

Table D reveals that the principals' evaluation of the availability of equipment, supplies, clerical time, and consultant time reflected a wide range of opinions about the sufficiency or insufficiency of the items. The principal comments indicated that supplies and equipment were insufficient at the onset of the program.

TABLE D

Principal Evaluation

		Supplies		Equipment		Clerical Time		Consultant Time	
Scale from 1 to 5		f	%	f	%	f	%	f	%
Insufficient	1	7	19	7	19	6	17	0	0
	2	4	11	5	14	5	14	4	11
	3	10	28	9	25	13	36	2	6
	4	4	11	2	6	5	14	12	33
Sufficient	5	11	31	13	36	7	19	18	50

Table D is based on Form E4C, Item 4.

4.27 Supplemental Data Based on Form E4C, E4D, E4E, and E4F.

Of the 36 principals, 30 evaluated the overall effectiveness of the program as "highly effective" and six evaluated it as "effective." Principal ratings indicated the program should be continued and expanded. Parents, pre-school teachers, and other teachers on the staff were positive about the program. Pupil and community attitudes toward the program were also highly positive. The attitude of other teachers not assigned to the Pre-School Program was positive.

4.30 Outcomes

Parents indicated that their children had made improvement. The questionnaire showed that 419 of the 420 parents felt the project should be continued.

Teacher ratings revealed that pupils made significant progress. Teachers also indicated satisfaction with the project and strongly recommended its continuation.

Principals endorsed the program with enthusiasm.

5.00 CONCLUSION

The Pre-School Program helped pupils develop listening and speaking skills, learn how to play and work together, and learn how to care for themselves and for their possessions in a group setting.

Evidence indicated that the program enhanced to a significant degree a pupil's potential for success in the first grade.

Parents and school staffs indicated satisfaction with the project and recommended its continuation.

6.00 RECOMMENDATIONS

Extend the in-service education program for teachers and administrators to provide understanding of the pre-school program.

Maintain selected cluster groups for purposes of longitudinal study as pupils progress through kindergarten and higher grades.

Recruit volunteer aides giving careful consideration to motivation and reliability factors.

Provisions should be made for designing and constructing classrooms and related facilities which would serve the pre-school programs more adequately.

Provide for the approval and funding of the program in sufficient time to permit the ordering and delivery of instructional materials prior to the beginning of the program.

ADDENDUM A

LOS ANGELES CITY SCHOOL DISTRICTS OFFICE OF RESEARCH AND DEVELOPMENT

ESEA Elementary Design #4, #10

PROJECT: Preschool II and IV

OBJECTIVES	OPERATIONALLY DEFINED DEPENDENT VARIABLES	ASSESSMENT DEVICES	COMMENTS
To increase children's speaking and listening skills prior to their entry into grade 1	Teacher rating of pupil pronunciation and oral expression	Rating Scale (E4A)	Comparison of pre-post data in 39 schools
To help children learn to play and work together in a social situation	Number of critical incidents	Rating Scale (E4A)	Pre - post comparison 39 schools
To teach children to care for themselves, their personal needs, and their possessions in a group setting	Teacher rating of pupil self-care and hygiene	Rating Scale (E4A)	Pre - post comparison 39 schools
To enhance pupil's opportunity for immediate success upon entering grade 1	Scores on PPVT and Psychomotor Parent Scores on Vineland	PPVT Psychomotor Development Test (E4B) Vineland Social Maturity Scale	Pre-post of random sample 6 schools Post only to parents in 6 schools
To provide pupils and parents with supplemental services	Parent rating of services rendered	Questionnaire (E4F)	Post survey of random sample of parents of 6 schools
To provide teachers with special materials, equipment, and in-service training	Staff ratings	Principal Evaluation (E4C) Teacher Evaluation (E4E) Consultant Evaluation (E4D)	Assessment of strengths and weaknesses of the program by principals, teachers and consultants

ADDENDUM B

LOS ANGELES CITY SCHOOL DISTRICT
OFFICE OF RESEARCH AND DEVELOPMENT

TEACHER RATING SCALE

(1-2) Pupil's Name ____________________, ____________________
Last First

(3) Sex: Male ________ Female __________ (4-7) School Code ________

(8-11) Age: Year ________ Month __________ Teacher's Name __________

(12-13) Pre ____________ Post __________

Check one of the five categories below for each statement.

	Not at all 1.	Very little 2.	A fair amount 3.	Quite a bit 4.	A great deal 5.	No opportunity to observe 6.
(14) Child recognizes major body parts.						
(15) Child responds verbally to questions during conversation.						
(16) Child responds non-verbally to questions during conversation.						
(17) Child pronounces words correctly.						
(18) Child asks questions which imply an understanding of what has been explained.						
(19) Child speaks freely and is able to verbalize a need.						
(20) Child uses words correctly in terms of their meaning.						
(21) Child is able to reach an amicable agreement with another child regarding shared use of equipment and materials.						
(22) Child takes active part in group activities.						E4A

		Not at all 1	Very little 2.	A fair amount 3.	Quite a bit 4.	A great deal 5.	No opportunity to observe 6.
(23)	Child is aware of space limitations, i.e., pouring liquid into vessels or loading blocks on a cart						
(24)	Child seeks out another activity following a peer conflict which is decided in favor of another child by the teacher.						
(25)	Child willingly acts upon teacher requests.						
(26)	Child initiates conversational exchanges with adults.						
(27)	Child requires no assistance in restroom activities.						
(28)	Child utilizes alternative approach when initial method of problem solving proves insufficient.						

Approved by: Robert J. Purdy, Associate Superintendent
Division of Elementary Education

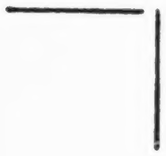
RETURN TO:

OFFICE OF RESEARCH AND DEVELOPMENT
at [illegible]
Elementary Project

BY: [illegible]

LOS ANGELES CITY SCHOOL DISTRICTS
DIVISION OF ELEMENTARY EDUCATION
Guidance and Counseling Section

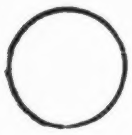
SCORING FOR EVALUATION FOR PSYCHOMOTOR DEVELOPMENT
Adapted from Rutgers Drawing Test*



Full Credit: 2 points if both lines are reproduced in a fairly accurate way. They can bend slightly.

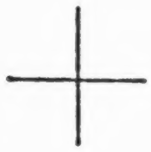
Half Credit: 1 point if only one line is reproduced fairly accurately.

No Score: If the child scribbles, or if he draws a vertical line in response to the horizontal line stimulus, or if he draws a horizontal line in response to the vertical line stimulus.



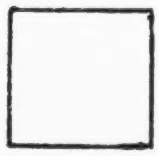
Full Credit: 2 points. Figure must be approximately round, have no angles; and lines must meet approximately at one point.

Half Credit: 1 point. Figure may not be round. It may be oval, etc., and it may contain some angles.



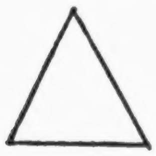
Full Credit: 2 points when both arms are of approximately equal length; are at right angles to each other; and bisect each other approximately. All lines must be firm and straight.

Half Credit: 1 point when figure resembles model, but when lines are not straight and when horizontal arm does not bisect vertical arm, but is above or below the midpoint of the vertical arm. Angles must be approximately right angles.



Full Credit: 2 points. Angles must be right angles; sides of figures must be approximately equal and parallel; and lines must be straight.

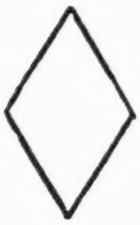
Half Credit: 1 point. Angles must be approximately right angles; sides may be unequal in length and lines may be somewhat irregular.



Full Credit: 2 points. Lines must be straight; sides must be equal but may be somewhat longer than the base and base must be parallel to horizontal lines on test paper.

Half Credit: 1 point. Lines may be somewhat irregular; sides need not be equal; one angle may be a right angle, or one angle may be somewhat rounded.

E4B



Full Credit: 2 points. Figure must be drawn in the approximate position of the model, the angles must be approximately equal as must the lower sides.

Half Credit: 1 point. Figure must be distinguishable from a square It must be in approximate position of the model; one set of angles may not be opposite each other; and upper and lower sides of figure may not be equal.

Derivation of Scoring Norms
Adapted from Rutgers Drawing Test

C.A.	Median
IV-0	2
IV-1	3
IV-2	4
IV-3	4
IV-4	5
IV-5	6
IV-6	7
IV-7	8
IV-8	8
IV-9	9
IV-10	9
IV-11	10
V-0	11
V-1	12
V-2	12

Report the child's score as the number of points successfully achieved.

If you want to relate this information to the teacher, you can make a comparison of the child's score with the median that corresponds to the chronological age. For example, if the child scores five points, his score would be comparable with the median score of a child IV-4.

---

* Taken from the Training School Bulletin, May, 1952, Volume 49, No. 3, by the Los Angeles County Superintendent of Schools, Division of Research and Guidance.

FG:rnc

LOS ANGELES CITY SCHOOL DISTRICTS
OFFICE OF RESEARCH AND DEVELOPMENT

PRINCIPAL EVALUATION OF THE PRE-SCHOOL PROGRAM

Principal________________________ School________________________

Date________________________

1. Please evaluate the overall effectiveness of the Pre-School Program which operated under your direction during the school year.

Highly Effective	Effective	Neutral	Ineffective	Highly Ineffective
___	___	___	___	___

2. Please estimate the impact of the program upon the pupils who participated.

Highly Effective	Effective	Neutral	Ineffective	Highly Ineffective
___	___	___	___	___

3. Please evaluate the attitude of the groups identified below toward the Pre-School Program.

Attitudes of Parents	Negative	___ 1	___ 2	___ 3	___ 4	___ 5	Positive
Attitudes of Pre-School Teachers	Negative	___ 1	___ 2	___ 3	___ 4	___ 5	Positive
Attitudes of other Teachers on staff	Negative	___ 1	___ 2	___ 3	___ 4	___ 5	Positive
Attitudes of Pupils	Negative	___ 1	___ 2	___ 3	___ 4	___ 5	Positive
Attitudes of the Community	Negative	___ 1	___ 2	___ 3	___ 4	___ 5	Positive

4. Please indicate the degree to which the following items were made available.

a. Allotment of equipment -	Insufficient	___ 1	___ 2	___ 3	___ 4	___ 5	Sufficient
b. Allotment of supplies -	Insufficient	___ 1	___ 2	___ 3	___ 4	___ 5	Sufficient
c. Clerical time -	Insufficient	___ 1	___ 2	___ 3	___ 4	___ 5	Sufficient
d. Consultant time -	Insufficient	___ 1	___ 2	___ 3	___ 4	___ 5	Sufficient

(over)

E4C

5. Please rate the extent to which you feel the services of the Consultant benefited the program.

Minimum ___ ___ ___ ___ ___ Maximum
1 2 3 4 5

6. What, if any, are the areas of greatest strength or of greatest value in the program?

____

____

____

____

7. Do you feel that the Pre-School Program should be continued at your school?

Yes___ No___

Comment______________________________

____

____

____

8. What recommendations do you have to improve the program?

____

____

____

____

Return to: Office of Research and Development at Emerson Manor, by May 6, 1966

Approved by: Robert J. Purdy, Associate Superintendent
Division of Elementary Education

LOS ANGELES CITY SCHOOL DISTRICTS
OFFICE OF RESEARCH AND DEVELOPMENT

CONSULTANT EVALUATION OF THE PRE-SCHOOL PROGRAM

Consultant______________________ School______________________

Date______________________

1. Please indicate on the 5 point scale below the extent to which you feel the program has been successful in meeting the needs of the pupils.

Minimum						Maximum
	1	2	3	4	5	

2. Please evaluate the effectiveness of the pre-service training as an aid to the teachers in the program.

Highly Effective	Effective	Neutral	Ineffective	Highly Ineffective
____	____	____	____	____

3. Please indicate the extent to which conference time with teachers was made available.

Insufficient						Sufficient
	1	2	3	4	5	

4. Please rate the extent to which you feel the travel involved within the geographical boundaries of your assignment affected the program.

Minimum						Maximum
	1	2	3	4	5	

5. What, if any, are the areas of greatest strength or of greatest value in the program?

____

____

____

6. What recommendations do you have to improve the program?

____

____

____

Return to: Office of Research and Development at Emerson Manor, by May 6, 1966

Approved by: Robert J. Purdy, Associate Superintendent
Division of Elementary Education

E40

LOS ANGELES CITY SCHOOL DISTRICTS
OFFICE OF RESEARCH AND DEVELOPMENT

TEACHER EVALUATION OF THE PRE-SCHOOL PROGRAM

School____________________

Morning__________________________ Afternoon________________ Date______

Please complete this questionnaire on the basis of your experience in the Pre-School Program during the 1965-1966 school year.

Individuals will not be identified in this study

1. Years of teaching experience

   Under 1 ___ 1-3 ___ 3-6 ___ 6-10 ___ Over 10 ___

2. Please indicate on a 5 point scale the extent to which you feel the program has been successful in meeting the needs of the pupils.

   Minimum ___ ___ ___ ___ ___ Maximum
   1 2 3 4 5

3. Please indicate the extent to which the following were made available to you.

		1	2	3	4	5	
a. Supplies	Insufficient	___	___	___	___	___	Sufficient
b. Equipment	Insufficient	___	___	___	___	___	Sufficient
c. Audio-visual materials	Insufficient	___	___	___	___	___	Sufficient
d. Consultant time	Insufficient	___	___	___	___	___	Sufficient

4. Please indicate how you feel the pupils were selected.
   This selection of pupils was:

   Inappropriate ___ ___ ___ ___ ___ Appropriate
   1 2 3 4 5

5. Please indicate the extent of parent contact.

   None at all ___ Very little ___ A fair amount ___ Quite a bit ___ A great deal ___

E4E

6. Please indicate the extent to which the consultant has provided assistance with:

Ideas:	Minimum	1	2	3	4	5	Maximum
Materials:	Minimum	1	2	3	4	5	Maximum
Techniques:	Minimum	1	2	3	4	5	Maximum
Environment:	Minimum	1	2	3	4	5	Maximum

7. What, if any, are the areas of greatest strength or of greatest value?

8. What recommendations, if any, do you have to improve the program?

RETURN TO:
OFFICE OF
RESEARCH AND DEVELOPMENT
at EMERSON MANOR

BY: May 6, 1966

Approved by: Robert J. Purdy
Associate Superintendent
Division of Elementary Education

LOS ANGELES CITY SCHOOL DISTRICTS

OFFICE OF RESEARCH AND DEVELOPMENT

May 6, 1966

Dear Parent:

The Los Angeles City Schools have provided classes for pre-schoolers. We are pleased that your child has had an opportunity to participate in the program.

Please help us to evaluate the program by answering the questions listed below. Please have your child return this letter to his teacher as soon as possible.

Thank you

1. Did your child show interest in the program? (Check one)

Almost none _____ Very little_____ A fair amount_____ Quite a bit____ A great deal_____

2. Did your child learn from the class? (Check one)

Not at all_____ Very little_____ A fair amount_____ Quite a bit____ A great deal_____

3. Should the classes be continued next year? (Check one)

No__________ Not sure________ Yes________

4. Would you like to write some comments or suggestions?________

____

____

E4F

Case Study V

EVALUATION OF
A PROGRAM FOR DEVELOPING SPEECH AND LANGUAGE SKILLS IN THE EDUCATIONALLY DEPRIVED CHILD THROUGH THE UTILIZATION OF THE SPECIALIZED TRAINING OF SPEECH THERAPISTS

MILWAUKEE PUBLIC SCHOOLS

INTRODUCTION

This project, organized and conducted by the Milwaukee Public Schools, Division of Curriculum and Instruction, Department of Special Education (Speech Therapy), is funded under Title I of the Elementary and Secondary Education Act of 1965 and is set up to serve children in primary classes from disadvantaged homes who are encountering an oral verbal language delay. The evaluation of the operation of the first semester's program was conducted by the Department of Psychological Services and Educational Research of the Milwaukee Public Schools.

The general purpose of the project is to increase the verbal and conceptual ability of educationally disadvantaged pupils by developing speech and oral language skills through the utilization of the specialized training of speech therapists.

Dates of Inception and Conclusion

The planning for this project began in October of 1965 before the research associate was employed and assigned to evaluate the project. The actual operation of the project began on February 3, 1966, and the portion of the project described in this report concluded at the end of the semester, June 10, 1966. The project is being continued during the current school year. The same person who directed the evaluation activities during the first semester is responsible for the evaluation of the current operation.

The report which follows describes specific activities of the research person assigned to the project and includes a description of the design of the research study, a description of the sample selection procedures, a discussion of the data collection plan and procedures, and a brief description of methods used to analyze the data collected and report results to staff personnel and other interested persons. There will be an emphasis on the role and function of the research associate in these activities.

Population Served by the Project

This project was carried on in seven elementary schools involving 17 classes with a total of 532 children, ages 6-8, with a grade level of P1 through P4. Five of the seven schools were located in predominantly Negro neighborhoods, and two were situated in predominantly Spanish-speaking communities.

RESPONSIBILITIES OF PROJECT PERSONNEL

Speech Therapists

Project therapists were not in the main concerned with the conventional type of speech problems, such as articulation and pathology. These problems remained the responsibility of the school's regularly assigned speech therapist. The primary responsibility of the project therapists was to work with pupils on an intensive basis. Their work included speech and language evaluation, preparation of materials and planning of activities for speech and language stimulation, and compilation of materials and development of techniques found to be especially suitable in working with this type of child.

Project Director

The function of the project director was to administer the project as to selection of schools, classes, and project staff personnel. In addition to supervision of project activity she had the responsibility for in-service orientation of the project staff, for the ordering of supplies, and for the writing of summary reports and budgets as required.

Research Associate

Basic functions of the research associate included designing a research and evaluation plan for the project, establishing procedures and a timetable for data collection, and collecting data and preparing evaluative instruments as required to be used by project personnel in the evaluation of the project. In addition, her responsibilities included maintaining a close working relationship with all project personnel (especially the director and four project therapists), the writing of an interim or status report in April 1966, continuous feedback of information as to the status of the evaluation plan, final analysis of data collected, and the writing of a summary report on the results of the analysis of data in the project. Copies of the final report were forwarded to state and federal authorities as required and to other interested persons as requested.

EVALUATION

Initial Planning of Evaluation Procedures

As the first step in a sequence of evaluation procedures, the research associate arranged a meeting with the project director to discuss the general aims and goals of the project and to become more familiar with specific objectives as well as operational procedures. The research associate studied all of the objectives and, when necessary, restructured the objectives and reworded them in terms of pupil behavior amenable to measurement.

After the research associate had critically examined the objectives, she met with the project director again to review the objectives and to suggest changes which might be made to improve the project. At this time the project director was encouraged to rank the objectives in order of importance, or at least to differentiate between primary and secondary objectives.

Provided with a list of objectives, demographic data, and the sequence of operational procedures the project personnel intended to follow, the research associate worked with the project director in preparing an evaluation design which would include procedures for collecting data regarding the population or a representative sample, tentatively selecting measuring instruments to be employed, a data collection plan, and a data collection schedule.

While preparing the evaluation design the research associate obtained samples of standardized tests which she felt would be appropriate and prepared rough copies of survey forms and questionnaires to be considered for use in collecting data. She also reviewed the research literature available concerning this type of project and consulted with other research personnel and curriculum staff members to obtain their reactions to the plan she was developing.

The Research Design

At another meeting with the project director and the project speech therapists, the research associate presented the following research design which was approved and followed.

It was agreed that a three-group experimental design, with a sequence of treatment groups and a continuing control group would be followed. Pupils were randomly assigned to groups and the groups randomly assigned to treatments. Two experimental groups labeled X_1 and X_2, comprising two-thirds of the total project population of 532 pupils, were given the treatment. Test results for these two groups could be analyzed separately or combined into one experimental group totaling 345 pupils, since all of them were at approximately the same grade level, approximately the same chronological age, and were randomly selected from the total project population. The remaining one-third of the total project population (187 pupils) served as a control group.

Basic data, such as pupil birthdates, sex, grade level, Pintner-Cunningham I.Q. scores, number of semesters in present school, other schools attended, and whether or not the child was receiving speech therapy from the regularly assigned speech therapist, were collected for each of the 532 pupils in the project by the speech therapists using forms supplied by the research department.

In order to assess pupil change, all classroom teachers and project therapists were asked to rank their pupils as to their observed oral language functioning in group situations at the beginning and at the end of the first of two time blocks and to rank them again at the beginning and at the end of the second time block.

The evaluation plan also called for one small group of pupils from each school (seven groups) to be selected at random to participate in a five-minute tape recorded session of their performance at the beginning and end of each of the two time blocks. Four qualified speech therapists not connected with the project were asked to rate the tapes from each time block using a nine-item rating scale prepared by the research associate.

Attendance records for all pupils in the experimental group were kept and forwarded to the research associate. The project therapists were asked to keep a log of instructional materials and teaching techniques which they found useful in working with the pupils in the experimental groups. They also made an evaluation of the project as a whole at its conclusion.

The Ammons Quick Test was selected to be administered to a total of 251 pupils in the experimental and control groups who were picked at random from the total population. The test was individually administered by qualified substitute teachers who were given special training in proper procedures by the research associate at a special in-service training session. Forms 1 and 2 of the Ammons Tests were used in February as a pre-test. Forms 2 and 3 of the same test were administered to the same pupils by the same examiners in late May as a post-test. A combination of two forms of the same test was used upon recommendation of the Coordinator of Psychological Services who had been consulted for an opinion about the applicability of this test. Using two forms of the same test increased the validity and reliability of the test results since the test is simple and brief.

The Data Collection Timetable

A data collection timetable was prepared in order to assist all persons involved in the project. The schedule prepared by the research associate is condensed below:

Date	Activity
February 9	The research associate randomly selects pupils to be pre-tested.
February 11	Three substitute teachers are trained to administer the Ammons Quick Test by the research associate in consultation with the Coordinator of Psychological Services.
February 14-18	1. Forms 1 and 2 of the Ammons Quick Test are administered to pupils selected at random. 2. Project therapists collect basic data on all pupils in the project. 3. Testers return all tests and materials to the Educational Research Department for scoring by February 18.
February 21-23 (beginning of first time block)	1. The pupils are ranked in X group by the therapists, as instructed by the research associate. 2. The pupils are ranked in the total class by classroom teacher. 3. The therapists make tapes (pre) for first time block (X). 4. Rankings are due at Educational Research Office on February 24.
March 25	The basic data from the Cumulative Record cards are due at the Educational Research Office.

Date	Activity
April 1-7 (end of first time block)	1. Ranking of pupils in X group by therapists (post) 2. Ranking of pupils in total class by teacher (post for first time block and pre for second time block) 3. Therapists make tapes (post) for first time block (X) 4. Rankings due at Educational Research Office on April 7, also attendance sheets for first time block
April 18-20 (beginning of second time block)	1. Ranking of pupils by therapists (pre) X_2 2. Tapes (pre) made by therapists for second time block 3. Rankings due at Educational Research Office on April 21
May 27-June 3 (end of second time block)	1. Post-test of randomly sampled pupils from the project population using the Ammons Quick Test, Forms 2 and 3. Tests and materials are to be returned to the Research Department for scoring and analysis by June 3. 2. Ranking of pupils by therapists (post) X_2 3. Ranking of pupils in the total class by teacher (post) for the second time block 4. Tapes (post) made by the therapists for the second time block 5. Rankings and attendance sheets for the second time block due at Educational Research Office on June 6
May 24 and June 6	Tapes scrambled (pre and post mixed) by the research associate and rated by speech therapists outside of the project

Date	Activity
June 8	1. Principal and therapist evaluations sent to the Educational Research Department for analysis 2. Therapists' logs of most valuable instructional aids and teaching techniques collected by the research associate

Analysis of Data

Several statistical procedures were used by the research associate in the analysis of both objective and subjective data. The percentage of attendance for the experimental group in each time block was computed. Mean ages, I.Q. scores, dropout rates, and transfer records were analyzed.

Spearman rank correlations between teacher and therapist rankings pre and post for each time block, teacher and teacher rankings pre and post for each time block, and therapist vs. therapist rankings pre and post for each time block were obtained.

The Wald-Wolfowitz Runs test was applied to the teacher vs. teacher rankings for the randomly selected classes to appraise the consistency of the teacher rankings.

A Sign test was used to measure change in the position of teachers' rankings.

The five-minute taped sessions of randomly selected groups seen by the therapists on an intensive basis which were made pre and post in each time block were rated by four licensed speech therapists not connected with the project.

The research associate applied a statistical test (Sign test) to assess changes in ratings.

The research associate made comparisons of the mean differences of the scores on the pre and post Ammons Quck Test using t tests of significance of differences.

At the close of the project therapists were given forms prepared by the research associate to evaluate the suitability of instructional aids. Positive and negative comments were tabulated and recorded by the research associate.

The therapists' logs of instructional materials which they found most helpful in working with the experimental group were reviewed by the research associate and forwarded to the Project Director. Lists of some of the most valuable teaching techniques used in working with the experimental group and which could be used by regular classroom teachers in working with culturally disadvantaged children were categorized by the research associate.

At the conclusion of the project the project therapists were given forms prepared by the research associate to evaluate the project, and the principals of the seven project schools were asked to rate the project as to how well it met six criteria. The research associate tabulated these ratings. The research associate discussed the results of the data analysis with project personnel. Recommendations for improvement of the project were developed in these discussions.

A thirty-four page report of the project evaluation was prepared by the research associate for distribution to state and federal authorities as required and to other persons on request. The preparation, reproduction, and distribution of several hundred copies of this report was supervised by the research associate.

Case Study VI

AN EVALUATION OF
THE MORE EFFECTIVE SCHOOLS PROGRAM
NEW YORK CITY PUBLIC SCHOOLS

ORGANIZATION AND CONTEXT FOR EVALUATION

Multiple Uses of Evaluation

The evaluation of the More Effective Schools (MES) Program is conducted by the Bureau of Educational Research of the Board of Education of the City of New York. This evaluation, however, enlists the cooperation and the participation of the administrators and the staff of the More Effective Schools Program in collection of information and data. As the data are analyzed and summarized there is a systematic feedback, or communication, of the findings to the personnel conducting the More Effective Schools Program. They use the information and data for supervisory purposes, including modifications of their practices. The data are used, also, in various formats and interpretations in making reports to state and federal agencies, to the Board of Education of New York City, and in relatively non-technical format to parents and citizens.

Some information and data frequently have supervisory uses in that they enable supervisors to help teachers do a better job of teaching individual pupils and class groups. Teachers and supervisors use the achievement test data, for example, to help diagnose individual pupils' strengths and weaknesses and apply remedial measures. The research uses of the information and data are represented by interim as well as final reports to various agencies and groups.

Fulfilling Legal Requirements

The interim and final reports fulfill the following legal requirements:

a. Legal requirements by the Superintendent of Schools and Board of Education for local school reports or evaluations

b. New York State Education Department and Federal, or USOE, for Title I of ESEA evaluations as required by Federal legislation

Context for Evaluation

The Office of Educational Research in New York City coordinates the activities of three research bureaus, namely, the Bureau of

Educational Research, the Bureau of Educational Program Research and Statistics, and the Bureau of Curriculum Research. Each of these bureaus is responsible for certain aspects of the evaluation of the More Effective Schools Program.

In addition, a contract has been negotiated with the Center for Urban Education in which a panel of experts and impartial observers made an assessment of the More Effective Schools Program.

Data Analysis and Information Reporting

The Bureau of Educational Research and the Office of Educational Research are mainly responsible for data analysis and information reporting. Data of a statistical nature frequently involve the use of computers and other modern methods of data processing. These were used in the More Effective Schools Program.

Support Needs

The methods used in evaluating the More Effective Schools Program have resulted in a good climate in the New York City School System. There is general agreement on the need for impartial and objective data collection, analysis, and interpretation. The Board of Education and the Superintendent of Schools have provided adequate budgetary provisions for professional, clerical, and consultative staff to conduct an appropriate evaluation of the More Effective Schools Program.

PROGRAM DESCRIPTION

Origin of the Program

In 1964 a Joint Planning Committee was formed to investigate possibilities whereby the New York City Board of Education could develop facilities which would conserve and utilize as fully as possible the human resources represented by the children in New York City Schools. The Committee was made up of representatives of the Superintendent of Schools, the United Federation of Teachers, and the Council of Supervisory Organizations. The Committee issued the Report of Joint Planning Committee for More Effective Schools to the Superintendent of Schools on May 15, 1964.

Some Major Goals of the Program

The report of May 15, 1964 set forth the philosophy underlying what was to be called the More Effective Schools Program: "There are too many children in our community who are growing up without the basic skills necessary for future success as citizens. We believe that these children, properly challenged and given the means for growth and learning, can make unprecedented academic and social progress. To meet this challenge a new design for education must be created." The design recommended by the Committee was one that would focus on the prevention of academic failure in the early years by starting education at the pre-kindergarten level and by organizing small classes to insure individual attention for each child's needs. Many teachers of special subjects and a clinical team for each school were to be provided. Classes were to be heterogeneous; that is, children of varied achievement levels in a given grade were to be placed in the same classroom. Intensive teacher training was to be part of a program which included as major educational strategies team teaching and nongraded instruction.

The schools selected for the program were to be located in socially disadvantaged areas of the city. They were all previously to have been Special Service Schools, a designation which signifies that low reading level and the percentage of free lunches and English language handicaps indicate more severe problems than are found in other schools in the New York City System.

Specific Goals of the Program:

1. School Organization

a. Pre-kindergarten classes: All schools have established pre-kindergarten classes for four-year-olds, and some of the schools have classes for three-year-olds.

b. Class size: There are a maximum of 15 students in pre-kindergarten classes, 20 in kindergarten, and 22 in grades 1 - 6.

c. Clusters: Classes are organized to form clusters, each of which consists of two classes at the pre-kindergarten level and three in all other grades. Each cluster has an extra teacher, the "cluster" teacher, who does not have a home class of her own but spends one period or more each day with each of the classes in her cluster.

d. Preparation period: All teachers have one preparation period per day. This is made possible by the presence of the cluster teacher.

e. Heterogeneous grouping: Classes are organized heterogeneously; that is, at each grade level, there are children of varying abilities in each class. Within the class, however, the teacher can group and regroup according to interest and ability.

2. Personnel

a. Administrative assistant: Each principal has an administrative assistant who handles many of the organizing and scheduling duties that previously occupied much of the time of the principal. This is intended to free the principal to expand his supervisory and person-to-person functions in the school.

b. Assistant principals: Each school has at least three assistant principals. Each assistant principal covers one of the following groupings: pre-kindergarten through grade 2, grades 3 and 4, and grades 5 and 6.

c. Pupil personnel team and other special services: Each school has a full-time team made up of three guidance counselors, one psychologist, one social worker, and one attendance teacher. Each school also has the services of a psychiatrist and a clinical speech teacher one day a week.

d. Other teaching positions (OTP's and special teachers): During the 1964-1965 and 1965-1966 school years, each school had a team of approximately seven teachers who were selected by the principal to best meet the needs of the school in the following areas: library, reading instruction, corrective reading, art, music, audio-visual, science, language resource, and health education. A speech improvement teacher was supplied full-time for each school by the Bureau of Speech Improvement. (This was in addition to the part-time clinical speech teacher.) These teachers were used for teacher training, demonstration, and team teaching. They also covered classes in order that each cluster team of teachers could plan together one period weekly.

3. Integration

a. Location of schools: When possible, integrated schools were chosen for the program.

b. Reverse open enrollment: White parents have been sending their children to four of the More Effective Schools which contain predominantly Negro or Puerto Rican children.

c. Community relations coordinator: There is one coordinator on the staff of each school whose duty is primarily to involve the community in active participation with the school. He does this by contact with the PTA and community religious and social organizations. The coordinator also conducts discussion groups for parents and courses for teachers.

4. Teaching Methods and Materials

a. Flexible grouping: Each class reflects a wide range of interests and abilities, since classes are not organized homogeneously according to ability. Teachers are expected, however, to group within the class. Often there will be two teachers (class plus cluster or OTP) within one class. This allows for a variety of small group and individual instruction. In addition, grouping may occasionally take place within the whole cluster when children in all three classes having a particular interest or problem are brought together for special work.

b. Team teaching: With the older children, classes in a cluster are sometimes brought together with one class teacher, cluster teacher, or OTP teacher teaching the lesson. The large group is then broken up into small discussion or activity groups, each one being led by one of the teachers. The teachers in a cluster plan together as a team and coordinate lessons and teaching materials.

c. Supplies and textbooks: Each More Effective School receives an extra allotment for supplies, textbooks, and visual and auditory aids. Special emphasis is placed on texts and other materials which stress urban backgrounds and deal with city children of varied racial and economic backgrounds.

d. Instructional emphasis: The goals of the program are many, but prime emphasis is placed on the improvement of language skills in general and reading ability in particular.

EVALUATION PLAN

Objectives: Their Clarification and Use in Evaluation

The objectives for use in evaluation were selected and defined jointly by the operations team and the evaluation team. These objectives are the basis of a longitudinal study beginning in September 1964 and continuing for the duration of the experimental program. Continuous communication and feedback from the evaluation to the operations team is characteristic of the program.

Outline of Evaluation of The More Effective Schools In New York City

OBJECTIVES	METHOD OF ASSESSMENT
Program of ME Schools	
- School Organization	A panel of nine nationally recognized educators visited, observed, and evaluated various aspects of the program and administration of the More Effective Schools.
- Personnel	
- Curriculum	
- Pupil Activities	
- Parent Attitudes	
Process Changes	
- Average Class Size	Official periodic reports
- Pupil-Teacher Ratio	Official periodic reports
- Costs of Instruction (per pupil)	Payroll, supplies, etc. expenses
- Pupil Attendance	Official periodic reports
- Pupil Mobility	Special study of admissions
- Teacher Mobility	Special study of transfers/discharges
- Pupil Ethnic Census	Annual ethnic census of pupils
Academic Skills	
- Reading	Longitudinal study-Metro. tests
- Reading	Experimental-control comparison
- Arithmetic	Longitudinal study-Metro. tests
- Language Skills (Pre-K-6)	Tests and teacher ratings

Reaction of Participants

Assistant Superintendents	Interview and questionnaire
Principals	Interview and questionnaire
Teachers	Interview and questionnaire
Parents	Interview and questionnaire

EVALUATION FINDINGS

Implementation of the More Effective Schools Program Objectives

It was concluded that the more Effective Schools Program as it operated was not vastly different from the program as it was envisioned. Some provisions became actualities in their entirety. These pertained to class size, heterogeneous ability grouping, teaching materials, cooperation with local colleges, audio-visual techniques, teacher specialists, staff recruitment, teacher preparation periods, and the use of community relations experts. The remaining provisions received only partial implementation. Only half of the schools were integrated; there were many classes for four-year olds, but few for three-year olds; all schools used team teaching, but only one used the nongraded bloc method; the pupil personnel team contained appropriate personnel for handling emotional and social problems, but did not include sufficient medical personnel for physical problems; some courses were offered to teachers and some scholarships were available, but financing did not come from the Board of Education; teachers did receive a daily preparation period but not complete relief from all non-teaching duties; the school plant was used fully during the school day and the summer months but not during the weekends. However, there were no recommendations that were not at least partially implemented.

Selected Statistics Describing the Program

Analysis of the data on the ethnic composition of pupil enrollment before and after the 21 schools were designated More Effective Schools shows that there was relatively little change in the proportion of Negro, Puerto Rican, and Other pupils on register before and after the schools became involved in the program. Ten of the schools could be considered integrated to a reasonable degree.

A study was made of the cost of instruction per pupil during the 1965-1966 school year in the 21 MES schools and the 9 control schools.

For this study the cost of instruction was considered to include both the salaries paid to pedagogical and non-pedagogical personnel and also the expenditures for school supplies and equipment. The data show that the cost of instruction per pupil in the 10 MES schools established in September 1964 was $859.38; the cost of the 11 MES schools established in September 1965 was $930.35. These amounts greatly exceeded the cost of instruction per pupil in other city elementary schools which was $433.86 for the 1964-1965 school year. Cost data for the 9 control schools showed that their instructional expenditures per pupil were one-half of the cost in the MES schools.

A study of pupil mobility in the 10 Old and 11 New MES schools showed that, in the Old schools, changes in mobility before and after the first year of the program were generally very small. In the second year of the program (1965-1966) 8 of the 10 Old MES schools showed declines in rate from the previous year. For the 11 MES schools established in September 1965, analysis of mobility trend data provides little information since these schools have not been in the program long enough to determine their effect upon pupil mobility.

Teacher mobility data for the period October 1965 through June 1966 were also analyzed. Only 2.7 percent of the teachers transferred from their MES school to non-MES schools and an additional .4 percent transferred to other MES schools in the city. In the MES schools all teachers were given the option of transferring at the end of the school year. In non-MES schools only 5 percent maximum can transfer each year. The teacher mobility rates for all reasons combined (maternity, sabbatical, transfer, etc.) in the MES and control schools were found to be respectively 6.2 and 6.4 percent.

The additional teaching positions assigned to the 21 MES schools brought average class size and pupil-teacher ratio in both years of the program well below the average ratios for city elementary schools in general. The average class size for all 21 MES schools as of October 1965 was 8.2 pupils less than the average for all other city elementary schools; pupil-teacher ratio was 10.8 pupils less.

A study of pupil attendance in the 21 MES schools showed that there was practically no change in attendance rates before and after the start of the program. As yet, there is no evidence that the MES program has had any appreciable effect on pupil attendance.

Achievement Test Results

A study was made of progress in reading and arithmetic by means of standardized tests for pupils participating in the More Effective Schools program. For pupils in the Old MES schools such progress was analyzed for one and two year periods; for pupils in the New MES schools progress was studied over a one year period.

Achievement test data were analyzed in relation to national norms in two ways. Grade scores attained were compared with the national norms applicable at initial and final testings, and the net change in pupil status in relation to the norm was noted. In general, for the three separate substudies in reading, the net change was favorable. That is to say, the grade scores were higher in relation to the norm at the final testing than they were for the initial test.

The second method of analyzing the standardized test results consisted of a comparison of the gains in grade score made between initial and final tests with the expected gains based upon the elapsed time between the testings. Here again the results for the three reading substudies were favorable in that, in most cases, the gains achieved exceeded the gains to be expected on the basis of national norms.

In a separate study of progress in first grade readings over a period of five school months in the Old and New MES schools, the findings showed that pupil gains exceeded the expected growth over the period studied.

An additional study compared reading growth in the 21 MES schools with selected control schools matched on the basis of ethnic composition of pupil register and third grade median reading grade score. The data showed that the grade groups in both the Old and New MES schools showed reading growth equal to or exceeding that for corresponding grade groups in the control schools, except in the case of the fourth grade in the Old MES schools.

Analysis of the data on pupil progress in arithmetic problem solving in the Old and New MES schools during the 1965-1966 school year produced findings which, in general, paralleled the results in reading achievement. Change in standing relative to the national norm improved over the experimental period for most grade groups. Similarly, the gains achieved usually exceeded expected gains postulated on the basis of elapsed time between initial and final tests.

There were some exceptions to the generally favorable results, but there were no instances of very poor achievement.

Language Skills Projects

An analysis of oral communication skills of pre-kindergarten and kindergarten children in the MES schools was conducted during the 1965-1966 school year. The data were obtained from an Inventory of Oral Communication completed by classroom teachers and composed of 35 items grouped under the two headings, Expressive Ability and Receptive Understanding. During the 1965-1966 school year the inventory was completed by classroom teachers for 2,670 pre-kindergarten and kindergarten children in the MES program, and the findings indicated that both the pre-kindergarten and kindergarten groups showed significant improvement from fall to spring of the school year as indicated by the proportion of the children falling within the two highest rating categories for each item.

As part of the MES program, the Bureau of Speech Improvement sends one speech teacher full-time to each MES school, and each teaches 20 different classes one period per week. The speech program's goals focused on such areas as listening habits, attitudes toward oral expression, oral language development, audibility and voice quality, and articulation and pronunciation. Some approaches used to achieve these program goals were creative dramatics, choral speaking, role playing, group discussions, exercises in correct phrasing, and word and sound pronunciation. In order to determine the existence of problems in such areas of speech as audibility, attitudes, rate of speech, vocabulary, foreign accent, and vocal quality, a speech checklist was completed by the speech teachers in the fall and spring of the 1965-1966 school year for a sample of the children in the program. Analysis of the data showed that the greatest number of children had problems with audibility and dialect in both the fall and spring. Children had the fewest problems with rate and hesitancy. The greatest improvement in the children was shown with respect to withdrawn attitude and audibility while the least improvement was shown in slow rate and uncooperative attitude.

Reactions of Administrators, Teachers, and Parents to the More Effective Schools

Analysis of the reactions of district superintendents, principals, teachers, and parents in the program to certain aspects of it which were obtained from questionnaire responses in the spring of 1966 showed that the district assistant superintendents felt that reduced class size and the establishment of pre-kindergarten classes were very effective results of the program. They also indicated that more individualized instruction was possible as a result of the reduced class size and additional personnel. The principals' reactions were very similar to those of the superintendents; this group also cited reduced class size and more opportunity for individualized instruction as the chief benefits of the program. Both the principals and superintendents agreed that the program should continue, though 50 percent of the responding principals proposed some modifications, most of them minor. The 900 teachers responding indicated the reduced class size and the daily preparation period as the most important features of the program in helping pupils learn. The parent reactions to the program generally echoed those of the administrators and teachers. This group cited improved reading as the outstanding benefit resulting from the program. Overall, for administrators, teachers, and parents, the most strongly lauded aspect of the program was the small class size.

The principal reservations expressed concerning the program centered on doubts as to the desirability of heterogeneous grouping, the large number of additional school personnel, and the need to meet the demands of community groups desiring the establishment of an MES program in their schools.

CONCLUDING REMARKS

The findings of the appraisal of the More Effective Schools Program are generally favorable. The objectives have been implemented to a reasonable and satisfactory degree, considering all factors. Class size and pupil-teacher ratios have been very favorable. Pupil and teacher mobility present no major problems. Pupil attendance presents no problems.

Standardized test results in reading and arithmetic show favorable gains in ability and skills by the MES pupils whether or not they are compared in growth with national norms or with a comparable control group of schools. Speech and oral communication data also revealed growth in the pupils.

The reaction of administrators, teachers, and parents to the MES program was definitely favorable. They favored reduced class size, individualized instruction, teacher preparation periods, pre-kindergarten classes, and personnel for improved services.

Analysis of costs has made it clear that the MES program requires considerable funding. On the basis of the evaluation as a whole, it would appear that the program needs to be kept essentially undiluted if it is to remain effective. If such elements as small class size are not retained, it is quite possible that the educational results will not be as favorable as this report has shown them to be.

Case Study VII

EVALUATION OF ESEA REMEDIAL AND CORRECTIVE PROJECT
SENIOR HIGH SCHOOL
OAKLAND PUBLIC SCHOOLS

ORGANIZATION OF ESEA TITLE I EVALUATION SERVICES

I. Description of General Research Department Services

A. The Director of Research is directly responsible to the Superintendent of Schools. The Research Department is composed of the Director, four Assistants in Research, fifteen school psychologists and psychometrists, four teachers on special assignment, and twelve classified staff members. The department works closely with various staff members in the elementary, secondary, special services divisions, and other district departments and offices.

B. The major responsibilities of the Research Department are as follows:

1. Individual Testing Program. Department personnel trained in psychological assessment techniques administer a variety of individual tests each year for pupil case study information as well as assisting in the identification of pupils for various special program (MR, gifted, etc.).

2. Group Testing Program. The Department coordinates and supervises the district group testing program and provides the following services to the district: purchase of all tests, distribution of tests, test scoring, summarization and analysis of test results, and consultant services on interpretation and use of tests.

3. Evaluation of Special Projects. Numerous studies are made each year to evaluate the effectiveness of various federal, state, district, and private foundation-funded programs. Examples are projects funded under ESEA Title I, EOA, VEA, NDEA, etc.

4. Administrative Research Studies. This function involves a variety of administrative studies including district surveys of class size, salary schedule surveys of other districts, and salary schedule development.

II. Research Department Involvement in ESEA Title I Project Development

A. Elementary and secondary committees developed suggestions for ESEA Title I programs. Representatives of the Research Department served on both committees.

B. Final proposal decisions were made by the Superintendent, Assistant Superintendents, Business Manager, and Directors of the following departments: Elementary Education, Secondary Education, Special Urban Educational Services, and Research.

C. Two Researach Department staff members wrote the final proposal application.

D. Because project development has expanded into many additional areas recently, the proposal development and writing function is now coordinated by the Department of Special Urban Educational Services. As proposals are being developed, the Director of Research or his representative does the following things:

   1. Meets regularly with the planning committee
   2. Provides data on pupil needs and characteristics
   3. Provides data for determination of schools in target areas

III. Research Department ESEA Title I Evaluation Functions

A. Assists in proposal development
B. Assists in defining project objectives
C. Provides staff for program evaluation. Present staff includes the following:

   1. One Assistant in Research.
   2. Two and one-half teachers on Special Assignment positions.
   3. Three clerical positions.

D. Develops the evaluation design. The design is developed and plans for implementation are carried out by an evaluation committee which meets for a full morning meeting every other week throughout the school year. Members of the evaluation committee are the certificated staff members listed in C above, the Director of Research, the Assistant in Research in charge of the district group testing program, and the Director of the Operating Program.

The development and implementation of the evaluation design involves the following:

1. Selection and development of appropriate tests and other data-gathering instruments
2. Development of a data-gathering schedule
3. Distribution of data-gathering devices
4. Provision for data storage and management, including a master file or data bank for project subjects
5. Provision for and conduct of appropriate data analysis procedures

E. Writes evaluation reports. This function involves the preparation of complete reports as well as needed abstracts.

F. Disseminates findings. This involves providing appropriate reports for school staffs, the district administration, the Board of Education, the local CAP Education Committee, State Department of Education, other school idstricts, and the press.

INTRODUCTION TO REMEDIAL AND CORRECTIVE PROJECT

Large numbers of students attending the target high school are performing two or more years below grade level in reading, written and oral language, and other skill areas. This is attested to by the test results which will be presented later in this report and by daily classroom performance. The academic problems of many of the students are reinforced in a cyclical fashion by poor attendance and marginal motivation for school work. These problems are particularly critical in light of the fact that many of the students are nearing the end of their formal education and therefore "must" possess the basic skills necessary for employment and otherwise productive lives.

The administration and staff of this school selected the English program, which included many interrelated areas in the language arts, as the focus of the ESEA remedial and corrective project. All 22 "B" and "C" English section students and staff members were designated to be involved in the program in varying degrees. Of the 320 students in grades 10, 11, and 12 who were served by the program, approximately 140 were selected to receive intensive assistance with attendance follow-up and individualized classroom instruction.

In order to provide a more individualized instructional program, major additions of staff were made to the department. The English Department Chairman was freed of in-class responsibilities in order to coordinate the program as a teacher assistant. Six reading teachers were also added to the English Department to work with the twelve regular teachers in all "B" and "C" sections. Their primary function was to provide small group instruction for the students who were most greatly in need of individualized attention. Since students were taken from existing class groups, this plan had the secondary effect of reducing class size in the regular classes.

The regular teacher and reading teachers worked together in selecting six to eight students from each class for the small group instruction program which emphasized motivational activities, word attack and comprehension skills, basic skills in self-expression, oral and written, and building improved self-confidence. Special cubicles were erected in rooms not having facilities for the second teacher. In some teaching teams the two teachers worked more or less exclusively with their respective groups. In other situations the composition of the groups was more fluid and/or the teachers exchanged students for various activities.

New materials including ditto masters for worksheets, reading "labs", new paperback literature books, tape recordings, and filmstrips were purchased under this program for utilization by both the reading teachers and the regular English teachers. In addition, tape recorders and tachistoscopically controlled filmstrip projectors were purchased. Although much of this supplementary material and equipment actually arrived and was employed in the program, it should be noted that some of the items mentioned above were not available until very late in the semester or after the school year was completed.

Teachers were provided clerical assistance for the preparation of study and drill sheets and other teacher-made materials. The clerk also worked closely with the teacher assistant in facilitating the testing program and other data collection, preparing reading lists, compiling bibliographies of professional references, and handling the general load of administrative details.

With added staff provisions two-thirds of the teachers in the program were able to have an added preparation period for the development of materials to meet the unique educational needs of the students. Innovation and experimentation were central themes of the program. It was felt that there was an urgent need to deviate from many of the traditional instructional units, materials, and techniques. Working together in meetings and respective teams the staff explored a wide variety of professional resources including textbooks, monographs, and periodicals. In their regular meetings, staff members also reported on materials and techniques which were felt to be effective. Much effort was made to locate materials which were of particular interest to minority students living in urban surroundings--those dealing with contemporary problems and situations. Two major in-service meetings, which featured consultants selected by the faculty, were held at the high school.

One unique feature of the language program was a school-wide "Better Speech Campaign." Classes examined the communication process, language patterns, articulation of sounds, and listened to various recorded language samples. Posters illustrating various agreed-upon objectives were prepared and displayed in the school. Various students compiled tallies of common errors in articulation and usage. Teachers were encouraged to devote twenty minutes daily to this month-long campaign.

Improvement of the educational progress of many of the students was integrally dependent on the simultaneous improvement of attendance patterns. The basic strategy employed here was that of closer communication with the parents of the students. Three parent aides worked with the teacher assistant, principal, and vice-principals on attendance problems. The parent aides met with this group to receive instruction as to their responsibilities and to offer comments or criticisms regarding the procedures suggested. Teachers submitted absence reports for all ESEA students.

Parent aides working with the teacher assistant and clerk typist selected approximately twenty names each morning--most of the names were those of students who were in the small groups. Parent aides worked three hours each day at home calling parents to ascertain why students were absent from school or from their English class. The following morning parent aides submitted typewritten reports of information received from parents. Appropriate action was instituted by the teacher assistant in cooperation with the regular attendance clerk, nurse, or vice-principal, if immediate action was necessary.

PROCEDURE

Subjects:

Students: As mentioned in the introductory section, 320 students enrolled in the 22 "B" and "C" section English classes were involved in the program. Although grades 10, 11, and 12 were represented in these sections, this report will deal primarily with data for grades 10 and 11. This decision is consistent with the overall evaluation design which includes plans for following the progress of students beyond this spring semester. Test data, then, are reported for two groups--(1) all tenth- and eleventh-graders in the program as a whole and (2) tenth- and eleventh graders who received the more intensive treatment of individualized instruction and attendance follow-up.

Faculty: Only English department faculty members were involved in this evaluation study.

Parents: A random selection of 40 parents of students involved in the intensive treatment program was designated as a pool of interview subjects. The pool was proportionately representative of grade level and pupil sex distributions. The plan was to interview a total of 25 parents from the basic and/or alternate lists within the pool. Within the time period available 19 interviews were completed.

Testing Instruments:

California Achievement Test, Junior High Level Battery: Four sections of CAT Junior High Level Battery were administered to students early in the program (pre: February 18-25; form X) and near the end (post: May 23-27; form W). There were approximately 11 weeks of school attendance possible during the pre-post -test interval. While grades 10-12 were tested at the beginning of the program to aid in the selection of small group attenders, only grades 10 and 11 were post-tested and only the results for the latter group are reported herein. The CAT yields a variety of sub-test scores which are reported in the tables in the FINDINGS section.

CAT SCORES

1. Math Vocabulary
2. Science Vocabulary
3. Social Studies
4. General Vocabulary
5. Total Reading Vocabulary (includes 1-4 above)

6. Following Directions
7. Reference Skills
8. Interpretations
9. Total Reading Comprehension (includes 6-8 above)

10. Capitalization
11. Punctuation
12. Word Usage
13. Total Mechanics of English (includes 10-12 above)

14. Spelling

STEP Essay Test: This is a free-writing activity. Students are given a standard topic which includes directions for an expository response. In the 4B form which was used in this pre-testing of the intensive treatment group, a brief account is given of an incident in which "Your school and its strongest rival have become involved in an argument over a close football game. . . . Write a letter to your principal suggesting ways in which these student councils may be able to restore a good relationship between the schools . ." This particular test was chosen since it dealt with a topic close to the experience of these students.

The test is unique in the sense that such factors as neatness receive weighting in the scoring. "Conventions," which includes such things as capitalization and punctuation, receives only a 20 per cent weighting. The main considerations here are "quality and thought" (50 percent) and "style" (30 percent).

Two substitute teachers, with training as English teachers, worked with the investigators in developing a consistent scoring rationale in line with the publisher's guidelines. Each scorer then completed an independent rating of each essay on a seven point scale. Where there was agreement between the ratings, that value was used. In cases in which different values were assigned, an average of the two ratings was computed. To accommodate for the "in between" values, for example, 2.5 or 3.5, interpolations were made on the publisher's percentile tables.

As noted above, data are available only for pre-testing and are reported at this time for descriptive purposes. Post-testing follow-up will be done during the spring of 1967, after a year's interval.

California Study Methods Survey: Sub-scales of this instrument cover the areas of "Attitudes Toward School," "Mechanics of Study," and "Planning and System." The instrument yields a score in each of these areas as well as a total "study methods" or study habits score.

Because the chief purpose for administering this instrument was not to sample achievement, an attempt was made to eliminate the necessity of possessing reading skills of a "sophisticated" vocabulary in order to respond to the items. Following consultation with the test author and publishers, permission was granted to modify various items through the substitution of common synonyms in the place of words which were identified by program personnel as "completely alien" to the majority of students in these schools. A second major deviation was the elimination of seven inappropriate items which did not apply to the curricular experiences of most Oakland Public Schools children. The test was presented to students in tape recorded form instead of the usual student-read booklets.

With the modifications of the instrument and because of the special characteristics of the study population (the published norms were based on a cross section of the general population), local "norms" will enable program personnel to rank students within this study population on various of the sub and total scores. They will also provide a baseline of data with which to compare results obtained in post-testing in the spring of 1967.

Data will not be reported herein; this aspect of the study is being described at the present time for purpose of noting the general status of overall data collection.

Program Staff Questionnaires:

Brief questionnaires, designed to elicit evaluative opinions regarding significant aspects of the program operation and effectiveness, were completed by English department faculty members in early June 1966. A copy of this questionnaire will be found in Appendix A. Responses were obtained from 15 of the 18 teachers involved in the program--an 83 percent return overall. Ten of the twelve regular English teachers' and four of the six reading teachers' responses are available for analysis.

Interview of Parents:

The plan for sampling parents' opinions regarding the ESEA program is briefly mentioned above. A complete description of the interviewing procedures is included in Report of Parent Interview Survey Regarding ESEA Program Services. A copy of the interview schedule used to sample parents' opinion of the ESEA program is included in Appendix B.

Sample Case Notes:

During the course of this remedial and corrective project teachers have kept brief logs of general classroom activities and the progress of individual cases. Samples of these self-report records of activities will be included in the following section.

FINDINGS

Results of California Achievement Tests:

Tables 1, 2, and 3 which follow present raw score means, standard deviations, and critical ratio comparisons of pre- and post-California Achievement Test results based on all tenth- and eleventh-graders for which both pre- and post-results were available. Approximately 65 percent of the subjects were tenth-graders; 35 percent eleventh-graders. It will be noted that the numbers of cases for sub scale score analysis are somewhat smaller than those used in the total scale scores (for example, see first four columns of Table 1). This is due to the fact that sub scale scores, such as Math Vocabulary and Science Vocabulary, were available at pre-test for only those students who used the IBM "1230" format answer sheets. It was not possible to obtain these sub scale score details from the IBM "805" format answer sheets which had to be used with approximately one half of the study subjects.

Composite means based on these sub scale scores approximate, but are not equal to, the means based on total scale scores. Since these values are very close, it is assumed that meaningful information regarding relative strengths and weakness in various sub skill areas may be obtained from these sub scale score analyses.

Grade score equivalents for mean values are tabled for the four battery sub tests but not for the sub scale values contributing thereto. Estimated values for the latter will be included in the test.

Examination of Table 1 (on the following page) will reveal that there were no significant changes evidenced in the mean raw score for "total reading vocabulary" between pre- and post-test administrations. It would appear, however, that there was significant growth in the sub skill area of "social studies vocabulary."

Results from this test indicate that these students are, on the average, more than three years "below grade level" in the area of reading vocabulary. These results do not reveal, however, whether the underlying deficiencies are related to word attack skill problems, conceptual limitations, or both.

Table 2 presents a more optimistic picture of progress in a second reading skill area--reading comprehension. Statistically significant increases in mean raw score values were observed for the total and sub scale. "Normal" month-for-month progress for the experimental period would have been slightly less than .3 grade equivalent units, or roughly three months. These students have, on the average, exceeded a normal rate of progress with a four month increase in scores. The fact that pre-test results were more than three and one-half years below a "normal progress" record adds even greater practical significance to these gains. In very roughly estimated terms these students prior to the ESEA program have achieved at a rate of less than 70 percent normal month-for-month growth:

$$\frac{\text{7.0 Actual Achievement}}{\text{10.5 "Normal" Grade Level Achievement}}$$

Therefore, even month-for-month growth may be interpreted as being greater than expected for this group.

TABLE I

Means, Standard Deviations, and Critical Ratio Comparisons of Pre, Post, and Difference Score Results from the California Achievement Test (Reading Vocabulary) - 10th and 11th Grade Students in Senior High School ESEA Program - Spring 1966

	Math Vocabulary	Science Vocabulary	Social Studies Vocabulary	General Vocabulary	Total Reading Vocabulary	
	Raw Score	Raw Score	Raw Scores	Raw Scores	Raw Scores	Grade Equiv.
Pre Test Feb. 1966 N	98	87	97	80	179	
Mean	8.57	6.82	6.30	7.05	30.67	7.1
σ	2.86	3.14	2.85	2.53	8.77	
Post Test May, 1966 N	98	87	97	80	179	
Mean	8.79	6.94	7.26	6.76	31.33	7.1
σ	3.19	3.67	3.36	3.65	10.92	
Pearson (Pre vs. Post)	.45	.45	.41	.43	.71	
Difference Scores (Post - Pre) N	98	87	97	80	179	
Mean	.172	.125	.939	-.272	.633	0
σ	3.22	3.62	3.40	3.424	7.731	
Critical Ratio	.53	.32	2.71	.71	1.09	
Level of Significance	$p > .05$ not signifi.	$p > .05$ not signifi.	$p < .01$ not signifi.	$p > .05$ not signifi.	$p > .05$ not significant	

TABLE 2

Means, Standard Deviations, and Critical Ratio Comparisions of Pre, Post, and Difference Score Results from the California Achievement Tests (Reading Comprehension) - 10th and 11th Grade Students in Senior High School ESEA Program - Spring, 1966.

	Following Directions	Reference Skills	Interpretations	Total Reading Comprehension	
	Raw Scores	Raw Scores	Raw Scores	Raw Scores	Grade Equiv.
Pre Test Feb. 1966					
N	98	97	93	207	
Mean	7.56	12.24	16.82	37.58	7.0
σ	2.52	3.83	6.14	10.74	
Post Test May, 1966					
N	98	97	93	207	
Mean	8.66	13.74	18.98	40.96	7.4
σ	3.01	4.51	7.21	12.68	
Pearson r (Pre vs. Post)	.50	.52	.51	.67	
Difference Scores (Post - Pre)					
N	98	97	93	207	
Mean	1.10	1.51	2.16	3.39	.4
σ	2.79	4.14	9.58	9.40	
Critical Ratio	3.89	3.63	2.17	5.17	
level of Significance	$p < .01$ Significant	$p < .01$ Significant	$p < .05$ Significant	$p < .01$ Significant	

Table 3 includes results for the remaining two tests of the CAT battery. In all instances the observed mean score gains were statistically significant. Again, the results equal or exceed what would be expected for "average" students during a comparable time period. Of particular interest in the analysis was the fact that the subjects' scores on the "Word Usage" sub test indicated a relatively more serious retardation in that area than in the other "Mechanics of English" sub scales. Approximated grade equivalent values for the respective means were as follows:

Capitalization	7.3
Punctuation	7.0
Word Usage	6.0

An analysis of the CAT results for approximately 50 of the cases receiving the more intensive remedial and corrective program is presented in Table 4. At pre-test, the mean status of the "intensive" group was from two to six grade placement units below that of the entire ESEA high school group. While mean raw score gains were made in all achievement areas tested, in only two areas,"Reading Comprehension" and "Mechanics of English", were these gains statistically significant. The "Reading Vocabulary" and "Spelling" mean raw score gains were smaller and accompanied by proportionately larger amounts of variability. The critical ratio values would indicate that mean gains of these magnitudes could be due to chance.

Special note should be given to the grade placement growth made in the area of reading comprehension. A half-year mean improvement over the 11 week period is particularly dramatic for a group of students identified by teachers as being those most in need of help.

Results of STEP ESSAY Tests

Results of the STEP Essay Test for 76 tenth- and eleventh-grade students involved in the "intensive" remedial classes are presented in Table 5 which follows. Since only pre-test results obtained in March 1966 are available at this time, the data are descriptive rather than evaluative.

TABLE 3

Means, Standard Deviation , and Critical Ratio Comparisons of Pre, Post, and Difference Score Results from California Achievement Tests (Mechanics of English and Spelling) - 10th and 11th Grade Students in Senior High School ESEA Program - Spring, 1966.

	Capitalization	Punctuation	Word Usage	Total - Mechanics of English		Spelling	
	Raw Scores	Raw Scores	Raw Scores	Raw Scores	Grade Equiv.	Raw Scores	Grade Equiv.
Pre Test Feb. 1966							
N	173	170	171	221		224	
Mean	19.63	13.48	23.36	55.44	6.8	15.44	7.4
σ	6.85	6.11	6.15	15.99		6.32	
Post Test May, 1966							
N	173	170	171	221		224	
Mean	20.82	15.03	24.66	58.84	7.1	16.60	7.8
σ	6.45	6.40	6.01	16.54		6.12	
Pearson r (Pre vs. Post)	.73	.71	.63	.81		.81	
Difference Scores (Post-Pre)							
N	173	170	171	221		224	
Mean	1.18	1.55	1.28	3.38	.3	1.12	.4
σ	4.92	4.78	5.24	10.13		3.83	
Critical Ratio	3.14	4.22	3.19	4.95		4.39	
Level of Significance	$p < .01$ Significant	$p < .01$ Significant	$p < .01$ Significant	$p < .01$ Significant		$p < .01$ Significant	

TABLE 4

Means, Standard Deviations and Critical Ratio Comparisons of Pre, Post, and Difference Score Results from the California Achievement Test (Reading and Language) - 10th and 11th Grade Students Receiving Intensive Program in Senior High School ESEA Program - Spring, 1966.

	Reading Vocabulary		Reading Comprehension		Mechanics of English		Spelling	
	Raw Scores	Grade Equiv.	Raw Scores	Grade Equiv.	Raw Scores	Grade Equiv.	Raw Scores	Grade Equiv.
Pre Test Feb. 1966								
N	40		47		52		53	
Mean	27.73	6.5	35.60	6.8	51.38	6.5	14.09	7.0
σ	10.53		12.85		17.38		7.13	
Post Test May, 1966								
N	40		47		52		53	
Mean	29.18	6.8	40.34	7.3	54.87	6.8	15.21	7.3
σ	12.47		13.10		18.06		6.68	
Difference Scores (Post-Pre)								
N	40		40		52		53	
Mean	1.45	.3	4.74	.5	3.48	.3	1.11	.3
σ	6.97		7.60		10.53		4.52	
Critical Ratio	1.29		4.23		2.37		1.76	
Level of Significance	$p > .05$ not significant		$p < .01$ significant		$p < .05$ significant		$p > .05$ not significant	

TABLE 5

Medians and Quartiles of STEP 4B Essay Test Results for 10th and 11th Graders Involved in ESEA Intensive Remedial Program at Senior High School - March, 1966.

ESSAY RATING	10th Grade			11th Grade		
	No. Cases	National Norms %ile	Median and Q's	No. Cases	National Norm %ile	Median and Q's
7.0						
6.5						
6.0						
5.5						
5.0						
4.5						
4.0	3	79		2	74	
3.5	1	69		1	62	
3.0	3	59		4	51	
2.5	2	47		6	40	Q_3 = 2.71
2.0	4	34	Q_3 = 1.94	5	29	Mdn = 2.05
1.5	7	25	Mdn = 1.20	5	20	
1.0	26	16	Q^1 = 0.97	7	11	Q_1 = 1.30
TOTALS		46 Students			30 Students	

Eleventh-graders tended to perform higher in terms of both norms, percentiles and raw score values. The percentile equivalent value for the tenth-grade median is approximately 20, for the eleventh-grade median 29.

The teachers who rated these essays prepared summaries of their observations. These reports are attached in Appendix C. One rater commented:

> "One of the first things I noticed in reading these essays was their vividness, awareness of life . . . Often, however, essays had to be rated down because of extensive problems of usage and mechanics"

These feelings were shared by the other rater who reported:

> "In general, students perform better 'stylistically' than they do in the wide realm of 'conventions' . . . . It is spelling and the entire punctuation - capitalization complex which seem to drag these essays down to various levels of sub-literacy"

Results from these essay tests and others administered on an optional basis apart from the evaluation study should provide useful "program information" to the instructional staff, that is, data regarding relative strengths and weaknesses which would have implications for the direction of the instructional program.

Program Staff Questionnaire Responses

Responses to the "Teachers' Questionnaire" were obtained from 15 of the 18 program staff members - - an 83 percent return. A copy of the questionnaire, with tabulated responses (number and percentages) and summarized comments, will be found in Appendix A.

Ninety-three percent of the teachers indicated that they felt the program was effective -- 13 percent "extremely effective"; 27 percent "very effective"; and 53 percent "effective." The reading teachers were more conservative in their evaluation. All four responded in the latter category. General support of the program team concept was expressed in comment to question 4, but teachers indicated that more time was needed to perfect the team approach.

Two thirds of the staff indicated that they had a four-period (instead of five-period) assignment and were unanimously agreed that the extra preparation time augmented their teaching effectiveness. Teachers' comments revealed that this time was most frequently used for materials development, planning, and completing "clerical - instruction" duties (correcting paperwork).

All but two teachers reported that they had made varying amounts of use of the new instructional materials provided under ESEA: eight mentioned different literature; nine mentioned the "skill builder" books and SRA Kits; two mentioned the special recordings of Negro Poetry. The reader is referred to the tabulated responses attached to the questionnaire in Appendix A for greater detail.

Ninety-three percent (14 of 15) had used the new audio visual equipment.

Teachers indicated a need for greater availability of various items although 73 percent rated the effectiveness of these teaching devices as "very good." Controlled readers were reported to be the most effective of those available. The reading teachers reported that they used the new materials and equipment somewhat more frequently than did the English teachers. They also tended to give higher effectiveness ratings.

Fourteen of the fifteen teachers noted improvements in the attendance of students in English classes as the result of the parent aide work with attendance problems: two reported "marked improvement"; twelve "some improvement." Sixty percent of the responders rated the clerical services as either "valuable" or "very valuable"; 26 percent rated them of little or no value; and 13 percent did not respond.

Table 6 presents the responses of teachers when asked to assess the magnitude of positive results which may have accrued as a result of the ESEA Program. "Improved motivation for learning" and " Increased self-confidence" were the two areas showing the greatest improvement. "Improvement in written language expression (content) " was ranked next in order.

TABLE 6

Responses of 15 Instructional Staff Members Involved in ESEA Remedial and Corrective Project at McClymonds High School, Spring, 1966

	Marked positive results	Moderate positive results	Limited positive results	No results evident	No Response
a. Improvement of oral language skills	(7%) 1	(47%) 7	(40%) 6	(7%) 1	-
b. Improvement in reading comprehension		(53%) 8	(33%) 5	(7%) 1	(7%) 1
c. Improvement in word attack skills		(53%) 8	(27%) 4	(13%) 2	(7%) 1
d. Improvement in written language expression (content)	(27%) 4	(33%) 5	(27%) 4	(7%) 1	(13%) 2
e. Improvement in motivation for learning	(47%) 7	(33%) 5	(20%) 3	-	-
f. Improved study habits (attitude and mechanics)	(7%) 1	(33%) 5	(33%) 5	(20%) 3	(7%) 1
g. Increased self-confidence	(47%) 7	(40%) 6	(7%) 1	(7%) 1	-

In response to a question asking teachers to cite the features of the ESEA program which had been most helpful, materials, equipment, class size, and the team approach were most frequently mentioned. The staff mentioned a scattered variety of problems which they had experienced during this spring's operation, some of which were the short duration of the program, too many tests, and the need for improved team communications (See Appendix A).

Results of Parent Interviews

Of the 19 parents of students involved in the intensive program, 53 percent indicated a familiarity with the new ESEA program. Thirty-seven percent had noticed changes in high school since the program's inception. It is surprising that only 42 percent acknowledged that their children were receiving special individual help. Of this group, however, all felt their children liked receiving the help. All of the parents felt that their children had benefited.

The fact that 68 percent of the parents reported that their children were more interested in school now than at the beginning of the program is congruent with the observations of teachers, noted above, but parents were generally not aware of the reasons for the increased interest.

Fifty-eight percent of the parents reported that their children do some reading at home, but only 8 percent noted any changes in their children's reading habits. Somewhat inconsistent with this is the fact that 30 percent of the parents felt that their children were reading more (in general) than in January. Approximately half of the parents noted improved spelling skills, more writing activities, and a greater amount of oral conversation.

Most of the parents interviewed (73 percent) were aware of the school's parent aide program, and 76 percent of this group felt that the aides were contributing significantly to the school's program. Onlythree parents, or 16 percent, indicated that they had attended a parent meeting at school. Seventy-five percent acknowledged, however, that they had been notified. Most felt that printed notices were the best method of informing parents of such meetings.

A smaller group of parents (37 percent) had heard of the Parent Advisory Committee. Those who had heard about this group had found out by mail, from friends and teachers, by phone, or in the newspaper. They had few suggestions as to what the Parent Advisory Group should be doing. Forty-two percent of the interviewees had heard about the Urban Renewal Program planned for the area but only one person had any suggestions as to what was needed in the area - - cleaning of the streets.

When these parents were asked about their "general impression of the job the Oakland Public Schools are doing in educating the children in your family?" the following responses were obtained:

Excellent - 2 (11%)
Good - 13 (68%)
Fair - 2 (11%)
Poor - 2 (11%)

None of the parents mentioned any specific problems in the school. Scattered suggestions were given for improving the educational program: more integration (2), more homework (1), smaller classes (1), increased reading programs (1), a full time psychologist at the site (1), a moratorium on suspensions (1), and before-school language and science courses (1).

Samples of Teacher and Aide Observations

The reading teachers, working with the teacher assistant, kept periodic case logs for students in the intensive small group program. These were used in planning and reviewing individual instructional programs over the course of the semester. The following entries will briefly illustrate the types of activities and strategies employed with a student in the program.

> 4-13-66
> The unit on mechanics was introduced today. He appeared amused when informed that his adeptness in usage and mechanics was not equivalent to his other skills. He readily agreed.
>
> 4-14-66
> Analyzed 10 sentences with aim of determining if E. could identify sentences. He has severe reading handicap. Became frustrated--eyes downcast, tears appeared.
>
> 4-19-66
> He requested my attention until 2:00--said he could do more work with me. Mechanics unit--worked on terminal marks of punctuation, and word attack skills.
>
> 4-28-66
> He asked if I knew anything about conjunctions. He had concept introduced by Miss Ba and felt he was getting a lesson devoted to this.

> 4-29-66
> Assignment given for homework and analyzed the next day. Student progress good. Confessed he had slapped the sister who wanted to help him.
>
> 5-3-66
> Analyzed relationship between parts of a compound sentence. Seems to have grasped relationship. Unsure of use of pronoun in second part of compound sentence. Frequently uses the wrong pronoun. Worked on pronoun and antecedent.
>
> 5-4-66
> Introduced term antecedent. Student wanted to know pronunciation and definition. Will not be deterred from obtaining material he desires. Wants to pursue a course he delineates.

At the end of the semester's work these teachers prepared synopses of experiences, impressions, and estimates of progress. The extract which follows is presented in illustration.

> C. came a long way in developing more positive attitudes during the past several months. The first month in the group he was very unsettled, did not participate in discussion, and often went his own way either mentally or physically. As time progressed, he began to participate more, his written work improved, and he began to develop self-control. . . . . .

Periodic reports were made by teachers to the teacher assistant. Brief synopses of activities and observations were noted and are illustrated in the following:

> Mrs. X
>
> Students discovered lack of library skills; class studied use of card catalog. Emphasis on group discussion and writing skills; increasing enthusiasm, participation, and confidence.
>
> Miss Y
>
> Interest shown by students in books they selected independently; interest in context clues, in new articles; interest aroused in discussing students' current needs.

Controlled reader seemed too "threatening" as used.

Mrs. Z

Work centered around the Better Speech Campaign--much interaction between students; improvement with use of tape recorder. Fact Recall most interesting activity of week.

Work in all classes on same materials. In preparation for the West Oakland Urban Renewal studies and visit by Mr. Kaplan of Urban Renewal, we read news magazine covering the Watts riots. My intent here was to get students to look for causes of what happened in Watts. I felt that if they were aware of possible causes in the Watts situation they might look with clearer eyes at West Oakland. As a follow up of readings we held a "news show" (K.G.O.) in which one student acted as moderator and each of the other students acted out the roles of Watts residents (mother of 3 sons, policeman, old resident of Watts. National Guardsman, unemployed).

Briefly describe one activity:

From the results of the Better Speech Campaign (my classes' tabulation of most common errors) and from students' compositions and speech, I am attempting to compile a set of exercises stressing common errors in sound, writing, and usage among students at this school. Each exercise is a series of related sentences. Students are asked to find errors and write in corrections. This requires very careful reading.

DISCUSSION

The overall impact of the available evaluative information is very encouraging, especially since the program has just been in operation for one semester. It is a monumental task to effect major program innovation, particularly at mid year.

The fact that students who were in this program were, on the average, three or more years below grade level suggests that most of them have experienced increasingly greater retardation with each successive grade level in their school careers. Instead of dropping farther behind, the group averages have generally "held their own, " so to speak, with month-for-month increases. It is also very significant to note that both teachers and parents have reported improved student interest and motivation. It is reasonable to speculate that there may be some relationship between the improved rate of progress and improved feelings toward school.

The lack of increase in reading vocabulary scores should be examined carefully. While this sub test is probably the most speeded test in the battery, the time factor was constant in both pre- and post-testing. The lack of a sufficiently developed reading vocabulary will undoubtedly prevent maximum development in other language arts skill areas.

During the course of the first semester's operation there have been opportunities for refinement of overall organization; new equipment and materials have arrived and have been explored; and teachers have had the benefit of experimenting with, or exploring, a variety of innovative instructional methods. With many of these developmental hurdles crossed it might well be anticipated that the progress of students next year may be even better--that they will come yet closer to average grade level performance.

If one accepts the responses of parents in the small interview sample as representative of a larger group of parents, the interview responses contain important implications for the area of home-school communication. It was somewhat surprising to discover that less than half of the parents of students in the "intensive" program were aware of their children's involvement in the formed ESEA program. The reasons for this are not clear. Perhaps students of this age do not inform their parents as fully as might be expected. Perhaps they have accepted the program as a "routine" type of experience which may speak well of the general atmosphere of rapport. Few parents had, however, come to the special meetings even though 75 percent of them were aware of them. The data do not reveal why. Since the parents have been getting printed notices, this vehicle of communication might profitably be used to more fully inform parents regarding program goals, services, and how parents can do their part in assuring maximum benefits to their children. There appears to be a positive feeling on the part of most parents regarding the job the Oakland Public Schools are doing. However, enlisting the active support of the parents continues to be a challenge.

SUMMARY

Approximately 320 students at McClymonds High School have been involved in a special remedial and corrective project focused in the area of the language arts--oral, read, and written. Limited auxiliary services in the form of clerical services and parent aide attendance workers were provided. The certificated teaching staff of the English department was augmented by the addition of six reading teachers. A variety of supplementary instructional materials and equipment was provided.

Achievement progress made by the total group of ESEA program students over the three-month test-retest interval generally equals or exceeds a month for month rate except in the area of reading vocabulary, in which no significant progress was evidenced. The mean achievement level of the "intensive" remedial group students was higher in all areas tested at the end of the program than at the beginning. In only two areas, "Reading Comprehension" and "Mechanics of English", were these gains statistically significant. While no "control" group data are available, it is readily apparent that month for month progress found in most test areas is greater than what would ordinarily be expected for a group which in ten or eleven years has fallen more than three years below grade level during their school careers.

The evaluative opinions of teachers and parents were clearly favorable. Both groups noted greatest improvements in the area of interest in and motivation for school. The fact that only 3 of the 19 parents interviewed reported they had attended special parent meetings and less than half were aware of their children's involvement in the "intensive" program emphasizes the need to continue efforts to improve home-school communication.

APPENDIX A

Teachers' Questionnaire - ESEA Program
At Senior High School, Spring 1966

A survey of faculty evaluative opinions and comments is being made in all ESEA schools now, at the end of the first semester of program operation. The results of these questionnaires will be incorporated in the report to the State Department of Education which will be prepared this summer.

Although some general types of information regarding teaching experiences and present teaching roles are being asked, there will be no effort to identify responders. This information will be used solely for grouping responses for analysis.

1. For how many years have you taught? ________

2. Is this your first year at this school? ________

3. Are you working primarily with the
   - (73%) 11 basic English classes (regular English Teacher)
   - (27%) 4 small group instruction program (ESEA Reading Teacher)

4. In your opinion, how effective has the present organization of having regular English teachers cooperatively working with reading teachers been in improving the language skills of students at McClymonds?
   - (13%) 2 Extremely effective
   - (27%) 4 Very effective
   - (53%) 8 Effective
   - (7%) 1 Not very effective
   - ☐ Not at all effective

   Comments: See attached tabulations of comments

5. How many class periods per day have you had this semester?
   - (67%) 10 4 periods
   - (20%) 3 5 periods
   - (13%) 2 No response

a. If 4 periods, do you feel the extra preparation period has increased your teaching effectiveness?

(100%) [10] Yes, definitely

☐ Yes, a little

☐ No

☐ Do not know

b. If 5 periods, do you feel that an extra preparation period would increase your teaching effectiveness?

(67%) [2] Yes, definitely

☐ Yes, a little

☐ No

(33%) [1] Do not know

Comments: (for Question 5) See attached tabulations of comments

6. To what extent have you made use of the supplementary instructional materials provided under the ESEA program?

(20%) [3] Used them extensively

(53%) [8] Used them Occasionally

(13%) [2] Used them infrequently

(13%) [2] Have not used them

a. Have quantities been sufficient? (53%) [8] Yes (20%) [3] No (27%) [4] No Response

b. Overall, how would you rate their effectiveness?

☐ Excellent (47%) [7] Very Good (27%) [4] Good (7%) [1] Fair ☐ Poor (20%) [3] No Response

c. Please describe 2 or 3 types of ESEA-provided materials which were effective, those which you would recommend for continued use.

See attached tabulations of comments

7. To what extent have you made use of the audio-visual equipment (Tach-X, Controlled Readers, tape recorders, etc.) provided under the ESEA program?

(40%)	(40%)	(13%)	(7%)
6 Used it extensively	6 Used it occasionally	2 Used it infrequently	1 Have not used it at all

a. Have quantities been sufficient?

(40%)	(53%)	(7%)
6 Yes	8 No	1 No response

b. Overall, how would you rate the effectiveness or value of the audio-visual equipment?

(33%)	(40%)	(7%)	(7%)	(7%)	(7%)
5 Excellent	6 Very good	1 Good	1 Fair	1 Poor	1 No response

c. Which type of equipment has proved most effective? ____________

See attached tabulations of comments

8. In your opinion, has the work of the parent-attendance workers had an effect on the attendance of students in your English classes?

(13%)	(80%)	(7%)
2 Yes, marked improvement	12 Yes, some improvement	1 Don't know

9. How valuable have the ESEA clerical services been to you in the preparation of materials, tests, worksheets, etc.?

(20%)	(40%)	(13%)	(13%)	(13%)
3 Very valuable	6 Valuable	2 Of little value	2 Of no value	2 No response

10. For each of the following areas please indicate your estimate of the results which were achieved with the majority of students with whom you worked in the ESEA program.

	Marked positive results		Moderate positive results		Limited positive results		No results evident		No Response	
a. Improvement of oral language skills	(7%)	1	(47%)	7	(40%)	6	(7%)	1		
b. Improvement in reading comprehension			(53%)	8	(33%)	5	(7%)	1	(7%)	1
c. Improvement in word attack skills			(53%)	8	(27%)	4	(13%)	2	(7%)	1
d. Improvement in written language expression (content)	(27%)	4	(33%)	5	(27%)	4	(7%)	1	(13%)	2
e. Improvement in motivation for learning	(47%)	7	(33%)	5	(20%)	3				
f. Improved study habits (attitude and mechanics)	(7%)	1	(33%)	5	(33%)	5	(20%)	3	(7%)	1
g. Increased self-confidence	(47%)	7	(40%)	6	(7%)	1	(7%)	1		

11. Please cite two or three features of this semester's ESEA program that helped you most to do effective work with the students.

See attached tabulations of comments

12. What problems were there in this semester's ESEA program which may have limited the effectiveness of your work with the students?

See attached tabulations of comments

Attachment to Teachers' Questionnaire - ESEA Program at McClymonds High School, Spring 1966

Question 4 - Summary of comments (frequency of mention)

Need to spend more time on basic skills - 2
Team approach good but takes time to perfect - 4
Team approach better with brighter students - 1
Small group particularly helpful - 2
Results will become more evident as program continues - 1
Need clearer statement of objectives - 2

Question 5 - Summary of comments (frequency of mention)

Need period because of extra forms, etc. - 1
Used time for lesson development, innovation - 4
Special class projects require much preparation - 2
Adequate preparation essential to good instruction - 2
Five period day is very demanding in terms of finding time for paper correcting - 4

Question 6 - Summary of comments (frequency of mention)

Tach - X (tachistoscopic filmstrip) - 1
Controlled reader - 2
Prepare study sheets (dittos) - 2
Roosevelt Grady - 1
New Worlds of Literature - 3
Recordings of Negro Poetry - 2
Sets of books, e.g., Native Son - 1
Oral Language Guide (McClymonds) - 1
Catcher in the Rye - 1
Reading skill booklets - 5
SRA Kits - 4
Plays for Modern Youth-1
Lord of the Flies - 1
Nectar in a Sieve - 1

Question 7 - Summary of comments (frequency of mention)

Tape recorder - 8
Controlled reader - 7
Phonograph - 2
Tach - X - 3
Motion Picture projector - 1

Question 11 - Summary of comments (frequency of mention)

Team approach - 3
Clerical Services - 1
Attendance workers (parent aides) - 2
Special equipment and materials - 5
Class size - 5
Freedom to experiment, innovate - 2
In-service activities, speakers, exchange of ideas - 5
Homogeneous (special needs) grouping - 1
Preparation time - 1

Question 12 - Summary of comments (frequency of mention)

Need for more planning time - 2
Too much standardized testing and paperwork - 2
Need more communication between team members - 2
Need for familiarity with more/better techniques - 1
Interruptions, need for continuity - 2
Midyear start was difficult - 2

APPENDIX B

May 27, 1966

Dear Parent:

Early in February, 1966, the new Elementary and Secondary Education Act (ESEA) program was begun in a number of the Oakland Public Schools. The purpose of the program has been to help students improve their reading, writing, and speaking abilities.

During the month of June, representatives of the Research Department of the Oakland Public Schools will be conducting an interview survey of many parents whose children have been receiving help under this new program. The purpose of these interviews will be to attempt to find out the opinions parents have about the ESEA program.

Within a few days an interviewer will be calling at your home. I hope that you will be able to take the time for the interview. Your views and opinions will be very helpful to us as we evaluate our present program and plan for the future.

Sincerely,

Alden W Badal

Alden W. Badal
Director of Research

Interview No. __________

This cover sheet used
for all levels of Parent
Interview Questionnaire

ESEA PARENT INTERVIEW

Interviewer______________________________Date Assigned_____________

Date Letter Sent________________________ ☐ No Letter

Resident's Address__________________________________Apt. Number___

Call No.	Date	Hour	Result of Call
1			
2			
3			

Time Interview Began________________Time Interview Ended______.

Comments:

FMM:ld
5/27/66

ESEA INTERVIEW QUESTIONNARE
(Senior High)

Hello. I'm ________________, a member of the survey team from the Research Department of the Oakland Public Schools. May I come in and talk to you?

We are talking to parents throughout the area to get a few of your ideas or opinions of the new Elementary, Secondary Education Act government program in the schools. It is also referred to as the ESEA Program.

1. You have a child ______________ in ______________ at ________________
(name) (level) (school)

Is that correct?

A Yes B No ☐

1a. IF YES: Since this is a survey requiring some questions I'm going to ask, I'll first say that any comments you make will be strictly confidential with no reference as to names or addresses.

IF NO: Oh, I'm very sorry. We are only interested in talking to parents in the area who have children in the Oakland Public Schools. Thank you very much anyway.

2. Are you familiar with the new program in the schools which began in February of this year?

	N	%		N	%	
A Yes	10	52.6	B No	9	47.4	☐

Total--19

2a. IF YES: Could you tell me some of the things you know or have heard about the program?

IF YES: CHECK OFF ITEMS IN QUESTION 3 AND IF ANY ARE OMITTED, ADD THEM WITH: In addition to those you mentioned--

IF NO: GO DIRECTLY TO NUMBER 3

3. MENTION ALL ITEMS IN YOUR OWN WORDS: Well, briefly, some of the changes are:

A More teachers_____(Hired)
B Parents for attendance work____(Hired)
C Reading specialists____(Added)
D New reading programs____(Started)
E New language programs___(Started)
F New supplies and equipment______(Started)

4. Now, I'm going to ask some questions about your feeling regarding ________________'s activities in __________________School.

5. But first, for how many years have any of your children attended _______________School?

	N	%		N	%	
A Less than 1 year	4	21.1	D 5 to 6 years	0		
B 1 to 2 years	9	47.4	E More than 6 years	5	26.3	☐
C 3 to 4 years	1	5.3				

Total--19

6. All right, I believe we have mentioned the changes in all the schools. Have you noticed any changes at _________School since last January?

	N	%		N	%	
A Yes	7	36.8	B No	12	63.2	☐

Total--19

6a. IF YES: What is your opinion of __________ School now compared to last January before you noticed these changes? Is it much higher, somewhat higher, about the same, or lower?

	N	%		N	
A Much higher	4	57.1	D Lower	0	
B Somewhat higher	3	42.9	E No opinion	0	☐
C About the same	0				

Total--7

7. What is your opinion about ______________'s ability to do the work in school now as compared to last fall? Would you say the work is much more easily understood now, somewhat more easily understood now, or a little more easily understood now, or hasn't there been any change?

		%	
A Much more easily understood now	5	26.3	
B Somewhat more easily understood now	6	31.6	
C A little more easily understood now	1	5.3	☐
D No change	7	36.8	

8. Has _____________ been receiving any extra reading help since January of this year?

	N	%
A Yes	8	42.1
B No	9	47.4
C Don't know	2	10.5
Total--19		

☐

8a. IF YES: Would you say (he) (she) likes it very much, likes it somewhat, likes it a little, or doesn't like it?

	N			
A Likes it very much	5	62.5	D Doesn't like it	0
B Likes it somewhat	3	37.5	E Don't know	0
C Likes it a little	0			
Total--8				

☐

8b. In your opinion, how helpful do you think this extra help has been? Would you say very helpful, somewhat helpful, a little helpful, or of no help?

	N	%	
A Very helpful	4	50.	D Of no help
B A little helpful	4	50.	E Don't know
C Somewhat helpful	0		
Total--8			

☐

9. Now from your observations of ___________, what are your feelings about (his) (her) interest in school now as compared to last January? Would you say there's much more interest now, somewhat more interest now, or a little more interest now, less interest now, or hasn't there been any change?

	N	%
A More interest now	8	42.1
B Somewhat more interest now	5	26.3
C A little more interest now	1	5.3
D No change	2	10.5
E Less interest now	3	15.8
F Don't know		
Total--19		

☐

9a. IF YOU HAVE A CHECK IN ANY BOX OTHER THAN No change: Well, what do you think has made this difference?

10. Now a few questions about ____________'s reading. How helpful do you feel this new government program is in helping (him) (her) to improve (his) (her) reading ability? Do you feel it's very helpful, somewhat helpful, of little help, or of no help?

		N	%			N	%	
A	Very helpful	6	33.3	D	Of no help	2	11.1	
B	Somewhat helpful	7	38.9	E	Don't know	2	11.1	☐
C	Of little help	1	5.6					

Total--18

10a. IF EITHER - Very helpful or Somewhat helpful - IS INDICATED: Concerning this improvement in reading, would you say there is much improvement, some improvement, or a little improvement as compared to last January?

		N	%	
A	Much improvement	4	33.3	
B	Some improvement	6	50.	
C	Little improvement	2	16.7	☐
D	Don't know	0		

Total--12

11. How much does ___________ read at home? Very much, some, a little, or practically none?

		N	%			N	%	
A	Very much	3	15.8	D	Practically none	6	31.6	
B	Some	8	42.1	E	Don't know	1	5.3	☐
C	A little	1	5.3					

Total--19

11a. IF READING AT HOME IS INDICATED: Off-hand, what are some of the things (he) (she) reads at home?

11b. IF READING AT HOME IS INDICATED: Have you noticed any changes in the kinds of material (he) (she) is reading as compared to last January?

		N	%	
A	Yes	1	8.3	
B	No	10	83.3	☐
C	Don't know	1	8.3	

Total--12

11c. IF YES: Well, what are some of the changes?

12. Compared to last fall, would you say (he) (she) reads much more now, reads somewhat more now, reads about the same amount, or reads less than last fall?

		N	%			N	%	
A	Reads much more now	3	15.8	D	Reads less	1	5.3	
B	Reads somewhat more now	3	15.8			2	10.5	☐
C	Reads about the same	10	52.6					

Total--19

13. Has there been any change in the amount of writing __________ is doing now as compared to last January?

A	Yes	8	42.1
B	No	11	57.9
C	Don't know	0	
	Total--19		

☐

13a. IF YES: Does (he) (she) write much more now, write somewhat more now, or write a little more now?

		N	%
A	Writes much more now	2	25.0
B	Writes somewhat more now	6	75.0
C	Writes a little more now	0	
D	Writes less now		
E	Don't know		
	Total--8		

☐

14. How does __________'s spelling ability compare now with (his) (her) ability last January? Would you say (he) (she)spells much better now, spells somewhat better now, spells a little better now, or there has been no change?

		N	%
A	Spells much better now	6	31.6
B	Spells somewhat better	4	21.1
C	Spells a little better	3	15.8
D	No change	6	31.6
E	Spells less well now	0	
	Total--19		

☐

15. Has there been any change in the things __________talks about at home now as compared to last January?

		N	%
A	Yes	3	15.8
B	No	13	68.4
C	Don't know	3	15.8
	Total--19		

☐

15a. IF YES: Well, what changes have you noticed?

16. As compared to last fall, does__________ talk much more now, talk somewhat more now, talk a little more now or there has been no change?

		N	%			N	%
A	Talks much more now	5	26.3	D	No change	6	31.6
B	Talks somewhat more now	4	21.1	E	Don't know	0	
C	Talks a little more now	4	21.1				
	Total--19						

☐

17. How about (his) (her) ability to pay attention to what others are saying as compared to last January? Do you think (he) (she) listens much better now, listens somewhat better now, listens a little better now, or hasn't there been any change?

		N	%	
A	Listens much better now	6	31.6	
B	Listens somewhat better now	6	31.6	
C	Listens a little better now	2	10.5	☐
D	No change	5	26.3	
E	Don't know	0		
	Total--19			

18. Now, some information about library books. Does ________ bring library books home?

		N	%	
A	Yes	5	26.3	
B	No	13	68.4	
C	Don't know	1	5.3	☐
	Total--19			

18a. IF YES: Which library does (he) (she) usually bring them from?

		N	%			N	%	
A	Public	2	28.6	C	Home	0		☐
B	School	5	71.4	D	Don't know	0		

Total--7 (Includes parent response to more than 1 item)

18b. IF YES: How does the number of books (he) (she) takes out now compare with the amount (he) (she) was taking out last fall? Would you say many more now, more now, a few more now, or hasn't there been any change?

		N	%			N	%	
A	Much more now	1	20.	D	No change	1	20.0	
B	Somewhat more now	1	20.	E	Don't know	0		☐
C	A little more now	2	40.					

Total--5

19. How much homework does ___________ bring home? Would you say very much, some, or a little?

		N	%			N	%	
A	Very much	4	21.1	D	None	3	15.8	
B	Some	8	42.1	E	Don't know	0		☐
C	A little	4	21.1					

Total--19

19a. Compared to last January does (he) (she)bring home much more now, somewhat more now, or a little more now?

A	Much more now	4	21.1	D	No change	10	52.6
B	Somewhat more now	4	21.1	E	Don't know	1	5.3
C	A little more now	0					

Total--19

☐

19b. Compared to last January how much time would you say (he) (she) spends on (his) (her) homework, much more time now, somewhat more time now, or a little more time now, or hasn't there been any change?

		N	%				
A	Spends much more time now	5	26.3	D	No change	6	31.6
B	Spends somewhat more time now	5	26.3	E	Don't know	1	5.3
C	Spends a little more time now	2	10.5				

Total--19

☐

20. Have you noticed any changes in ___________'s attitude toward (his) (her) homework since last January?

		N	%
A	Yes	4	21.1
B	No	13	68.4
C	Don't know	2	10.5

Total--19

☐

20a. IF YES: Well, what changes?

21. Perhaps you have already told me this, but did you know that Oakland Public Schools have employed parents and community residents to assist teachers to work with student attendance?

		N	%			N	%
A	Yes	13	72.2	B	No	5	27.8

Total--18

☐

22. Have you talked to any of these teacher assistants?

		N	%			N	%
A	Yes	7	36.8	B	No	12	63.2

Total--19

☐

22a. IF YES: What is your opinion concerning the value of this help for teachers and students? Would you say this help is very valuable, somewhat valuable, a little valuable, or of no value?

		N	%			N	%
A	Very valuable	5	71.4	D	Of no value	1	14.3
B	Somewhat valuable	1	14.3	E	Don't know		
C	A little valuable	0					

Total--7

23. Have you been able to attend any of the parent meetings held at McClymonds High School since the first of the year?

	N	%		N	%	
A Yes	3	15.8	B No	16	84.2	☐

Total--19

23a. IF YES: About how many meetings have you attended?

	N	%		N	%	
A 0			D 5-6			
B 1-2	(2)	66.7	E More than 6	(1)	33.3	
C 3-4			F Don't know			☐

Total--3

23b. IF YES: How did you find out about these meetings?

	N	%	
A Printed notice	1	33.3	☐
B Student told parent	0		☐
C Told by neighbor or friend	1	33.3	☐
D Other (Specify)	1	33.3	☐

Total--3

23c. IF NO: Were you ever notified or told about parent meetings?

	N	%		N	%	
A Yes	12	75.0	B No	4	25.0	☐

Total--16

24. What is the best way for the school to notify parents about meetings for parents?

	N	%	
A Printed notices	13	65.0	
B Telephone calls	6	30.0	☐
C Other (Specify)	1	5.0	

Total--20

25. Have you ever heard about the Parent Advisory Committee at McClymonds High School?

	N	%		N	%	
A Yes	7	36.8	B No	12	63.2	☐

Total--19

25a. IF YES: How did you find out about it?

25b. IF NO: GO ON TO QUESTION NUMBER 26.

26. What do you feel this Parent Advisory Committee should be doing.. .. how can this group help to improve the educational program at McClymonds?

	N	%	
A Don't know	14	82.4	
B No response	3	17.6	☐

Total--17

27. Have you heard about the special plan which is being developed for the McClymonds area? It's called the McClymonds Problems and Urban Renewal Proposal.

	N	%
A Yes	8	42.1
B No	11	57.9

Total--19

☐

27a. IF YES: How good do you think this plan is?

	N	%
A Excellent	4	50.
B Very good	4	50.
C Satisfactory	0	
D Fair	0	
E Poor	0	

Total--8

☐

Comments:

28. What do you feel would be the best way for the school to keep parents informed about important school matters?

Comments:

29. What is your general impression of the job the Oakland Public Schools are doing in educating the children in your family? Would you say excellent, good, fair, or poor?

	N	%
A Excellent	2	10.5
B Good	13	68.4
C Fair	2	10.5
D Poor	2	10.5
E Don't know	0	

Total--19

☐

30. Now, as a parent, have you had any problems in working with the school?

31. What suggestions would you make to the Oakland Public Schools to improve the educational program for your child?

APPENDIX C

Impressions of a STEP Essay Rater

ESEA Program - (Rater A)

Very generally speaking, the essays seem to reflect a picture of at least average minds (seemingly in a majority of cases anyway) struggling to cope with the accepted standards of "style" and "conventions." In general, students perform better "stylistically" than they do in the wide realm of "conventions." In a great number of cases, the writing flows beautifully, in terms of syntactical patterns. It is spelling and the entire punctuation-capitalization complex which seems to drag most of these essays down to various levels of sub-literacy. Also, a great many display sentence structure breakdowns--chaotic, distorted, jumbled "word hashes" which, ironically, probably do not appear orally, or somehow do not disrupt communication when the speaker is talking to members of his own non-standard English-speaking group. Somehow, the process of writing may slow down or distort the mental flow of utterances most people of any group are capable of. Perhaps future remedial English teaching could center around narrowing the gap between spoken and written English, especially when instructing minority group children or children from homes in which another language is spoken. Perhaps one small mode of attack would be reducing initially the importance attached to punctuation and spelling, thus "freeing" academically-limited students from the overwhelming task before them. The essays seem to reflect a syntactical and mental paralysis induced by an awful awareness of "literate" conventions--punctuation, usage, etc. So many booklets are profuse with erasures and scribbles--and blank spaces.

Impressions of a STEP Essay Rater

ESEA Program - (Rater B)

One of the first things I noticed in reading these essays was their vividness, awareness of and enjoyment of life. They contrasted tremendously with the bland, aloof style in which the comparison essays were written. Many times I felt frustrated when I had to rate an essay "2" or "3" when the insights in it showed a more mature outlook than the "4" or "6" comparison essays.

The essays were usually more lively, more direct, and made much more interesting reading than the comparison essays of the same level. Often, however, essays had to be rated down because of extensive problems of usage and mechanics. These factors added to confused or nonexistent organization seemed to make impossible the ordered, sequential, logical development which we ordinarily associate with good thinking and good writing.

Even the best essays were much less abstract than the example essays. Where a level 3 example essay might speak in general about the behavior of American families, a test essay most often showed the writer's own family. These essays made exceedingly interesting reading because of the use of concrete detail. One has the feeling that, once trained, many of these students would do very well in fictional creative writing.

Indeed, every tenth or fifteenth essay seemed to reveal a really active, imaginative, observant, energetic personality. The use of sensual imagery often made ideas vivid that more mechanically correct writers made dull and boring. Also I found dialect usage to be sometimes direct and forceful. Traditional teachers who see only usage or grammar "errors" in students' speech and writing may well be overlooking and failing to encourage communications skills which even the "poorest" students have.

Observations

Teachers must revise traditional ways of looking at student writing. The most important use of writing is to communicate ideas and observations. If nagging about technicalities destroys the students' interest in communicating, writing will be only an empty exercise. It might also be wise to look at other aspects and characteristics of schools which tend to limit or destroy students' desire to communicate ideas and observations.

Case Study VIII

EVALUATION OF ESEA TITLE I PROJECTS
PHILADELPHIA PUBLIC SCHOOLS

CONTRACTING WITH AN OUTSIDE AGENCY FOR EVALUATION

The volume of research undertaken through Federal program grants is such that many school districts, both large and small, find themselves faced with the problem of evaluating a large number of projects with insufficient staff and facilities.

The Philadelphia Public Schools embarked on 29 projects under Title I in the spring of 1966. The total professional and clerical staff of the Division of Research was insufficient to do many of the things that it felt should be done. It therefore seemed obvious that attempting evaluation of the Title I program would result in failure both for the evaluation and the ongoing programs. The solution seemed to lie in contracting for the required evaluation.

The first step was to list the agencies which were qualified to perform the type of service required. Each of the suggested agencies was then carefully studied to arrive at an informed judgment as to its capability for undertaking the proposed contract. The following questions were considered:

1. To what extent is the agency experienced in project evaluation?
2. Has the agency demonstrated ability to develop evaluation instruments as needed?
3. Does the agency have an adequate staff for handling the number of projects involved?
4. What is the machine and clerical capacity of the agency?
5. Does the agency have an adequate margin of capability above present commitments?
6. Can the agency assign a sufficient number of staff members to the contract to provide for simultaneous evaluation of projects?
7. How responsive would the agency be to emergencies?
8. What are the qualifications of the agency's staff in terms of training and experience?
9. What would be the probable cost elements involved in pricing the contract?

The number of agencies under consideration was not very great, and several of them were quickly eliminated because of a marked inadequacy in one or another of the aspects considered important. As the survey of agencies proceeded, one or two agencies were concentrated on. The final result of this process was the selection of the Franklin Institute Research Laboratories to perform the evaluations required under Title I.

The Institute has long had a national reputation for research in social and scientific areas, and it has performed evaluations similar to the one required for the Federal government and other agencies. Staff members are very satisfactorily qualified in terms of training, experience, and variety of capabilities. The Institute is a non-profit organization which tends to cut costs. It is well able to construct suitable instruments to meet needs that cannot be filled by materials already available. One of its greatest assets, which could not be matched by any other agency, was its proximity to the administrative offices of the school district. This made possible frequent and spur-of-the-moment conferences between the evaluation staff of the Institute and the personnel of the Division of Research. Another strong advantage was the Institute's printing facilities, which eliminated undue delays in the production of materials.

When the choice of agency was approved by the Board of Education and the contract was let, the Director of Research arranged to establish a working relationship between the staff of the Institute and that of the school district. In a joint meeting of the directors of all projects and the Institute's evaluation staff, the relationships between the two staffs were delineated and the Federal requirement for evaluation was made clear. One person in the Division of Research was designated to work closely with the Institute's staff so that the Division of Research would retain primary responsibility for the evaluation, with the Institute acting as its agent.

The persons assigned by the Institute to the responsibility for a specific aspect of the Title I evaluation met with the project staff of the school district as soon as posssible. The projects were thoroughly discussed so that the evaluation staff had a clear idea of the objectives of each project, of the methods and materials to be employed, and of the problems being faced by the project staff.

The Franklin Institute Research Laboratories developed 17 evaluation instruments used for the first time in Philadelphia's Title I program. The Institute staff was immediately available for "trouble shooting" as material shortages or problems of administration developed. All reports required by the State and Federal governments were prepared by the Institute and channeled through the administrative offices of the school district.

In Philadelphia the contracting of project evaluation to the Franklin Institute Research Laboratories was a successful solution to the problem of insufficient research capacity. This initial experience with Title I was an introductory phase which it is anticipated will be profitable in further Title I activities. The Philadelphia evaluation contract has again been awarded to the the Institute, and it is hoped that a full year of Title I will show more adequate evaluation and more valuable outcomes than were possible in the first brief experiences in the spring of 1966.

Although every research division would, undoubtedly, prefer to conduct its own evaluation of all experimentation and innovation within its district, it is questionable whether research divisions can or should provide themselves with the staff and equipment to bear the load of current Federal programs. To do so might result in considerable excess capacity if at some time the Federal programs should be dropped. In the present emergency, and perhaps even for future years, it would seem that the contracting of evaluation to responsible agencies should have a definite place in the thinking of those responsible for educational research.

EVALUATION BY THE FRANKLIN INSTITUTE RESEARCH LABORATORIES [1]

PROGRAM-EVALUATION APPROACH

The Franklin Institute Research Laboratories approach to the problem of evaluating programs for the disadvantaged is based on one inescapable fact: few tests have norms for this population. Therefore, we have attempted to develop unique tools which are applicable to the disadvantaged and the teacher who must teach this child; and which are at least partly tailored to the uniqueness of the particular school system and program being evaluated.

---

[1] The following pages are reprinted, with slight editorial change from Evaluation for Educational Effectiveness, pp. 2-5. Permission to use these pages was granted by the Franklin Institute Research Laboratories.

We attempt to build into each instrument some items which can be applied generally, and some which have reference to more localized conditions. Such a mix can be accomplished, we feel, only if the test developer spends time on-site to gain some insight into the local situation. Most programs clearly cannot be evaluated using tests already on the market. And even if positive results are obtained, severe questions of validity remain.

Because few standardized instruments are available, we approach the evaluation task without preconceived ideas of the correct "test" form to use. We may use checklists, or multiple-choice items, or Likert-type scales, or work-sample procedures. Multidimensional and multivariate methods are also under development. The measures presented here demonstrate this diversity.

The Educational Evaluation Team presently is developing a set of tests directed toward the disadvantaged child which use measurement techniques devised in our previous efforts. (For example, see the last measure presented in group IV here, the "Most-Like Questionnaire" -a multidimensional method of measuring self-concept.)

All these techniques are in the experimental stage. When they are available, data concerning the reliability and validity of these techniques will be reported to interested individuals. However, we hope it is apparent that the instruments are unique approaches to the problems of educational evaluation.

The FIRL approach does not discount traditional testing procedures with nationally standardized tests when they are applicable; in fact, more than half of our program may consist of such instruments. With disadvantaged children, however, results must be interpreted very carefully. For example, children may show improvement on an arithmetic test not because of improved facility in arithmetic, but simply because they can read the test items better. In addition because children may improve in one subject at the expense of lowered performance in another subject, we normally recommend across-the-board testing in all major subject areas.

MEASURING INSTRUMENTS

Educational-evaluation measuring instruments developed or modified by the FIRL team are of five types:

I. Achievement measures

II. Teacher-activity measures

III. Teacher-attitude measures

IV. Student-attitude measures

V. Other measures

Each measure (or part of a measure) is described briefly below.

I. Achievement Measures

A. Mathematical Knowledge (Form B) (selected pages). This revision of the Glennon Mathematical Knowledge Test is directed to teachers.

B. AAAS Science Test (Form B) (selected pages). This open-ended test has been revised into a multiple-choice format and is given to teachers.

C. Art Questionnaire for Elementary School. This test consists of questions of knowledge and of attitude toward art in about equal proportions.

D. Vocational-Training Questionnaire. This questionnaire asks questions pertinent to knowledge about jobs and job opportunities.

II. Teacher-Activity Measures

A. Counselor Checklist. Counselors indicate the amount of time spent working at various activities.

B. Counselor-Aide Checklist. Counselor Aides indicate the amount of time spent working at various activities.

C. Music-Teacher Checklist. Grade teachers indicate the amount of time they spend teaching 27 activities, 10 of which are musical.

D. Art Inventory. Art and grade teachers indicate whether they can teach various art experiences and whether they have taught them during the preceding year.

E. Academically Able Teachers' Questionnaire (selected pages). Teachers are asked how often they have used various techniques in extra classes. This also includes some attitudinal items and some questions of information.

F. Kindergarten Teachers and Kindergarten Aides Questionnaire (selected pages). Kindergarten Teachers and Aides indicate how much time they spend on various activities.

G. Teacher and Teacher-Aide Questionnaire (selected pages). Teacher and aides indicate how much time they spend on various activities.

H. Art Questionnaire for Elementary School. See item IC.

III. Teacher-Attitude Measures

A. Attitude Survey. This Mathematical Attitude Survey is given to the teachers who took the Mathematical Knowledge Test (see sample IA).

B. Title I Survey (selected pages). Given to teachers in all Title I programs, this survey assesses feelings about Title I programs, disadvantaged students, and related concepts.

C. Teacher Survey. Given to teachers in Title I programs, this survey attempts to uncover the teacher's concepts of the child.

IV. Student-Attitude Measures

A. Semantic Differential for Remedial Readers (selected pages). This test is given to students in reading programs on a pre-post basis to measure attitudes toward school, reading, teachers, self, and related items.

B. Student Questionnaire (vocational training) (selected pages). This measure is applied to vocational-training students to determine why they came to the program, how much they liked it, and what they thought of it.

C. Academically Able Questionnaire for Students (selected pages). This questionnaire was designed like the student Questionnaire used in the vocational training program. In addition to the items listed above, this asks about specific areas of accomplishment (such as grammar and speech) and about job attitudes.

D. Music-Attitude Test. Students are asked which of pairs of activities they prefer. Musical activities are compared to nonmusical activities.

E. Most-Like Questionnaire (selected pages). This questionnaire is given to students to evaluate self-image. Several forms must be used in a balanced design. One out of five pages of one of six counterbalanced forms is shown; within the six forms, all possible triadic combinations of twelve concepts are given.

V. Other Measures

A. School-Community Coordinator Program, Parents' Questionnaire. This questionnaire was given to a sampling of parents to measure reactions to the coordinator program.

B. School-Community Coordinator Program, Coordinators' Questionnaire. This questionnaire was given to coordinators to determine extent and type of coordinating activities.

In addition to the tests listed previously, the following standardized tests were used:

Stanford Achievement Test - Intermediate I and II, and Advanced

Test of Academic Progress

School and College Ability Test (SCAT)

What I Like To Do

Academic Promise Test

Differential Aptitude Test (DAT)

7

Kuder Occupational Interest

School Mathematics Study Group (SMSG) Test-Forms A and B

Sequential Tests of Educational Progress (STEP) Reading Forms 1A and 1B

STEP Social Studies Forms 1A and 1B

STEP Tests 2 through 4 (24 forms)

Gates Reading Survey